Why the Gene Will Not Return*

Author(s): Elisabeth A. Lloyd

Source: *Philosophy of Science*, Vol. 72, No. 2 (April 2005), pp. 287-310

Published by: The University of Chicago Press on behalf of the Philosophy of Science Association

Stable URL: https://www.jstor.org/stable/10.1086/432425

Why the Gene Will Not Return*

Elisabeth A. Lloyd†‡

I argue that four of the fundamental claims of those calling themselves 'genic pluralists'—Philip Kitcher, Kim Sterelny, and Ken Waters—are defective. First, they claim that once genic selectionism is recognized, the units of selection problems will be dissolved. Second, Sterelny and Kitcher claim that there are no targets of selection (interactors). Third, Sterelny, Kitcher, and Waters claim that they have a concept of genic causation that allows them to give independent genic causal accounts of all selection processes. I argue that each one of these claims is either false or misleading. Moreover, the challenge that arises from the availability of genic causal accounts, namely, the inability to choose on rational grounds among genic and higher-level accounts, is unsupported.

1. Surprising Announcements. From the 1960s through the early 1990s, substantive and far-reaching changes in population genetics occurred. The debates concerned the units of selection and changed the nature and scope of evolutionary explanations, turning attention from genotypic models to models of structured populations. Furthermore, the relation of a trait to its environment was explored in both models and empirical investigations.[1] E.g., Maynard Smith's early theoretical work on group selection provoked theoretical debate, while Michael Wade and colleagues produced empirical evidence for the efficacy of group selection as an evolutionary component. Wade's (1978) paper, in particular, criticized the ill-conceived, oversold,

*Received September 2004; revised February 2005.

†To contact the author, please write to: Department of History and Philosophy of Science, 1011 East Third Street, Goodbody Hall 130, Indiana University, Bloomington, IN 47405; e-mail: ealloyd@indiana.edu.

‡My ideas in this paper have benefited greatly from interactions with Stephen Downes, Melinda Fagan, Stephen Crowley, James Griesemer, Richard Lewontin, Ken Waters, Elliott Sober, Jay Odenbaugh, David Hull, Patrick Forber, Peter Godfrey-Smith, Gordon McOuat, Michael Dietrich, Rasmus Winther, Michael Weisberg, Roberta Millstein, Rob Cummins, Jordi Cat, Kent Van Cleave, two anonymous referees from *Philosophy of Science*, and the Biology Studies Reading Group at Indiana University.

1. Although Goodnight and Stevens (1997) convincingly trace these conflicts of focus back to Sewall Wright, I do not follow the rest of their analysis here.

Philosophy of Science, 72 (April 2005) pp. 287–310. 0031-8248/2005/7202-0001$10.00

and rigid theoretical requirements for the efficacy of group selection in nature. The result was a blossoming in a corner of population biology—both in theoretical genetics and in empirical studies—of investigations into the possibility of higher-level interactors such as those that had been suggested in the house mouse by Lewontin and Dunn 1960 and Lewontin 1962.[2] Various attempts had been made to expand population genetics beyond the organismically-focused efforts of the pre-1960s, but the work of the 1980s is distinguished by its achievement of consensus about many fundamentals of hierarchical population genetics and by its empirical substantiation of these ideas. Philosophers of biology soon joined the revolution, attempting to sort out the various disputes and definitions, analyzing the variety of evolutionary models, focusing in on what they saw as the key issue of contention: How could the various methods and principles be used for sorting out the key 'units of selection' (the objects directly involved in selection processes determining the success of genes), thus engaging in the 'interactor debates'.

In the midst of this scientific abundance, some philosophers, calling themselves 'genic pluralists', arose and attacked the new population genetics, its models, and reinterpreted its empirical findings.

In 1990, Philip Kitcher, Kim Sterelny, and Ken Waters (henceforth 'KSW') announced that no one had to concern themselves with the units of selection problem anymore: "Once the possibility of many, equally adequate, representations of evolutionary processes has been recognized, philosophers and biologists can turn their attention to more serious projects than that of quibbling about the real unit of selection" (1990, 161). Yet, the debates continue. Concerns about units of selection in biology and philosophy have not been settled by claims about representational equivalence.

I argue that three of the chief claims of the genic pluralists are defective, and that their overall argument is unsupported. First, there is the above-quoted remark about philosophers and biologists who quibble over the units of selection. Second, Sterelny and Kitcher (henceforth 'S&K') also claim that there are no targets of selection (interactors). Third, KSW claim that they have a concept of genic causation that gives independent genic causal accounts of all selection processes. I argue that each of these claims is either false or misleading. My approach to this literature is unusual; it doesn't take up the proffered notion of genic cause directly, nor does it defend another view of cause. Rather, my approach is indirect, proceeding through a fuller understanding of the role of interactors in selection theory. KSW's claim, that debate over units of selection becomes superfluous

2. These investigations are reviewed in Section 4. For detailed analysis of structured population models, see Lloyd [1988] 1994.

once the existence of equivalent (read: genic) representations has been established, is shown to have things exactly backwards: Problems of units of selection must be overcome in order to generate the adequate genic-level theory that they take as one of their "many, equally adequate, representations" (KSW 1990, 161). Debates over interactors, therefore, are not 'pseudoproblems', and this fact ultimately has fatal consequences for their claims that there are independent genic causal stories for any case of selection (KSW 1990, 161).

The burden of proof lies with the pluralists to show that any such genic-level causal accounts exist at all. I will not defend any particular hierarchical view in the paper, as these have been defended elsewhere.[3]

2. The Genic Pluralist Challenge. 'Pluralist genic selectionists' (such as S&K) believe that "there are often alternative, equally adequate representations of selection processes and that for any selection process, there is a maximally adequate representation which attributes *causal efficacy* to genic properties" (1988, 358; emphasis added). This sort of pluralism is peculiarly weak: It is simply an equivalence condition. Not only that, the arguments, as given by S&K, entail genic reductionism; an ironic twist, given that pluralism is usually an antireductionist position.[4]

More importantly, the peculiar sort of pluralism offered by S&K stands in contrast with philosophically significant, stronger forms of pluralism. There are three such forms. First, one in which each level of description is understood as indispensable and independent, either locally or globally, for describing a given phenomenon (Dupré 1993). This type of pluralism offers an important type of unity. Each level of description is insufficient, and connecting models are needed for the purposes at hand; this type of pluralism can be used to support a global disunity of science claim. A second important, and stronger, form of pluralism claims that either some or all the different descriptions are jointly relevant to the phenomenon under consideration; this is especially important when the relevant information cuts across different levels or different kinds of description (Darden and Maull 1977; Dupré 1993; Cat 1998, 2000, 2004). A third significant type of pluralism concerns unity and pluralism at the level of criteria, independent of any specific theory or application. An example would be multiple causal criteria, wherein criteria of causality above any specific theory contribute a new dimension to the question of unity and pluralism independent of the issue of the connection among theoretical facts (Cat

3. See Lewontin 1970; Wade 1978; Hull 1980; Wimsatt 1980; Sober 1984; Lloyd [1988] 1994; Brandon 1990; Williams 1992; Sober and Wilson 1998; Gould 2002, Section 4.

4. Waters' pluralism, though, is not reductionist, although it does share other weaknesses with S&K's pluralism.

2000, 2004). S&K's pluralism (further details of which are worked out in Section 5) exemplifies the close relationship between pluralism as equivalence and pluralism as reduction, both of which are in tension with the general treatment of pluralism in the wider literature, as is made clear in the conclusion.

KSW see themselves as attacking "some biologists" who "think . . . there is a unique account that will identify *the level* of selection" (1990, 159; emphasis added). This question of levels is the question of identifying the interactors for that process. But, they proclaim, "We believe that asking about the real unit [level] of selection is an exercise in muddled metaphysics" (KSW 1990, 159). This is because, according to KSW, the gene itself can always be construed as the interactor in a selection process—the entity that directly interacts with its environment such that replication is differential—once 'environment' is construed in terms of the allelic perspective (Waters 1991, 554, 571; S&K 1988, 339, 341, 348). Thus, "[hierarchical selectionists] err . . . in claiming that selection processes must be described in a particular way, and their error involves them in positing entities, 'targets of selection', that do not exist" (S&K 1988, 359).[5]

Thus, the pluralists' claim that there is a causally adequate, general evolutionary theory purely at the genic level—one that does not require any appeal to higher-level causal interactors. This is not so. By helping themselves to the necessary higher-level (interactor) information, the pluralists make it appear that the hard-won methods for obtaining and incorporating such higher-level causal information (which is represented in the mathematical structure of the models as a whole) are irrelevant; or that debates over these methods have been resolved or overcome by applying the genic approach. They have not. One of the primary claims of my paper is that, contrary to its proponents' claims, the genic account does not give us a theory independent of individuating causal interactions at various levels of the biological hierarchy, nor does it solve or dissolve the problem of how to individuate those very interactions.

3. The Basics. In Lloyd 1992, 2001, I analyzed *the* units of selection problem into four distinct questions, and argued that much confusion had arisen because participants in the debates were arguing at cross-purposes.

The four questions I delineated as distinct 'units of selection' questions were the replicator question, the interactor question, the beneficiary question, and the manifestor of adaptation question. The replicator question concerns which entity passes on its structure directly in replication (usu-

5. This may be a spot where Waters disagrees with S&K, since Waters does think of selection as a force acting on a target.

ally, but not only, genes) (Dawkins 1982; Hull 1980; but see Griesemer 2000). I shall expand on the second, interactor question, below, since it is the focus of this paper. The beneficiary question involves which entity benefits, in the long term, from the evolution by selection process, while the question of which entity manifests adaptations plays a central role in determining which entities have 'engineering' adaptations at a given level of organization (see Lloyd 2001; Williams 1966). Using this framework, I analyzed many of the major positions in the units of selection controversies, and concluded that numerous players were mixing and matching the various questions in developing their requirements for what it takes to be a real unit of selection.

One of the easy cases for my 2001 paper to sort out was the debate between Dawkins and the group selectionists. Dawkins makes clear that by a 'unit of selection', he means a replicator, not a vehicle (interactor).[6] Since he categorizes not only organisms but also groups as vehicles, groups cannot, by definition, be units of selection (1982, 115). In contrast, the genic pluralists have developed arguments, driven by new approaches to interactors,[7] which pit genic selectionists directly against group selectionists.[8] While numerous authors since 1988 have attacked these genetic pluralist views (originating in their present form with Waters 1986), none of these criticisms has been particularly successful in convincing readers exactly what is wrong with the position.[9]

I argue that the basic problem lies in the pluralists' understandings of the role of interactors in models. Their genic models are explicitly derived from causal models involving higher-level interactors, as I review in Section 5. Any genic causal account is thus derivative from an interactor causal account, and is not independent at the genic level, since it incorporates these higher-level causes. This result undermines their claim to have established independent causal genic selection models at all. Before presenting the details of this argument, let me return to the crucial interactor question.

4. Interactors. In its traditional guise, the interactor question asks, What units are being directly selected in a process of natural selection? An interactor may be at any level of biological organization, including group,

6. The distinction between vehicles and interactors is clarified in Section 4.

7. But see Williams 1966 and Maynard Smith 1987 as underdeveloped precursors.

8. These are arguments that I did not examine in my 2001 paper, although Lloyd [1988] 1994, 133–143 concerns them directly.

9. For example, Sober 1990; Sober and Wilson 1994; Shanahan 1997; Sober and Wilson 1998; Glymour 1999; Van der Steen and Van den Berg 1999; Stanford 2001; Glennan 2002; Wilson 2003.

kin-group, organism, chromosome, gene, or even parts of genes. The interplay between an interactor and its environment is mediated by traits that affect the interactor's effects on genic success. Some portion of the expected fitness of the interactor is directly correlated with the value of the trait in question. Finally, the expected fitness of the interactor is commonly expressed in terms of fitness parameters, that is, in terms of the fitness of replicators; hence, interactor success is most often reflected in and counted through, replicator success.[10]

At what levels of biological organization do interactions occur that make a difference to replicator success? There are a number of ways to study this question, including modeling, experimentation, and fieldwork,[11] and various methods have been proposed over the years for identifying interactors.[12] These have been used primarily in theoretical and philosophical discussions. At the same time, population biologists (sometimes in interaction with philosophers) have developed a wide variety of technical definitions.[13] There has been much argumentation over which approach is best.[14] The issue as it stands is undecided, although various research groups using different (but closely related) approaches, each with their various strengths, have made advances.

The emphasis in biological discussions of interactors is on getting the statistical and causal information that will make the model empirically adequate to the phenomena. For example, Heisler and Damuth's popular contextual analysis approach (Multilevel Selection I) to discovering and isolating interactors has three goals: the measurement of relationships between characters and fitness; the location of the *level* of biological organization at which these relationships occur; and the evaluation of *causal*

10. The term 'interactor' is David Hull's; it was designed to make up for shortcomings in Dawkins' term 'vehicle'. Specifically, 'vehicle' was meant by Dawkins to refer to the developmental consequences of replicators (not exclusively, but usually, genes). Problems with the vehicle idea occur however when genes themselves are the entities interacting with the environment directly, as in meiotic drive or segregation distortion. Using Hull's terminology, genes would, in these cases, be called 'interactors', whereas there is no place in Dawkins' hierarchies for them, despite the fact that he emphasizes these 'outlaws' from the organismic perspective (Hull 1988, 28; Kawata 1987).

11. See Lloyd [1988] 1994 and Sober and Wilson 1998 for relevant literature reviews.

12. This includes the approaches of Lewontin (1970); Sober (1984); Brandon (1982, 1990); Wimsatt (1980, 1981); Lloyd ([1988] 1994); Sober and Wilson (1994, 1998); and Glennan (2002), among others.

13. For example, Price (1972); Lande and Arnold (1982); Arnold and Wade (1984); Arnold and Fristrup (1982); Wade (1985); Heisler and Damuth (1987); Damuth and Heisler (1988); Lewontin and Dunn (1960); Lewontin (1962); Mayo and Gilinsky (1987); Nunney (1985); and Sober and Wilson (1998), among others.

14. For one superior recent critical evaluation, see Okasha 2004.

models of selection that are proposed on the basis of prior research (Heisler and Damuth 1987, 594–595). The struggles in the literature involve the best ways of doing so. Some of the empirical work[15] involves demonstrating the efficacy of various hierarchical models in accurately modeling an empirical system (Goodnight and Stevens 1997). Thus, constructing models with information at the higher-level is at the heart of the empirical and causal explanatory uses of hierarchical models. Dugatkin and Reeve emphasize explanatory (1994, 123–124) and research (1994, 126–129) successes of hierarchical models as against 'broad individual' models. Sober and Wilson, using the trait group approach to modeling hierarchical selection, review many case studies in which higher-level information makes the difference between building empirically, explanatorily, and causally adequate and inadequate models (1998). Griesemer and Wade (1988) discuss nontrait-group hierarchical models.[16] Despite the variety of models and the disagreements regarding the success of particular models, the take-home lesson here is clear. Empirical adequacy, explanatory sufficiency, and research promise are all evaluated to determine the appropriateness of interactor models at distinct levels of the biological hierarchy.

In Hull's analysis, selection must be understood as two distinct *causal* subprocesses: replication and interaction. In the genic selectionism of KSW (derived, as Waters acknowledges, from Williams 1966),[17] the same processes are involved; it's just that both take place at the genic level. The fundamental claim here is that genes are the ultimate interactors. Whenever other structures can claim to be interactors, genes can claim to be interactors, too, and the genic claims are more 'general' and 'unified'.

I now examine how the pluralist genic selectionists manage to formulate genes as interactors.

15. For example, house mouse (Lewontin and Dunn), Tribolium (Wade), insects (Colwell), ponds (Wilson), social hymenoptera (Dugatkin and Reeve), crop plants (Griffing), hens (Craig and Muir).

16. See discussion of the differences from trait group models in Wade 1978, 1985; Lloyd [1988] 1994; Goodnight and Stevens 1997.

17. Note that Williams does not adopt either Dawkins' later 'expansion' of his views, nor Waters' interpretation of his genic line of thought. Williams is committed to using the hierarchical notion of interactor: "Natural selection must always act on physical entities (interactors) that vary in aptitude for reproduction . . . interactors can be selected at levels from molecules to ecosystems, and there has been helpful recent progress on this levels-of-selection question" (1992, 38). Although, interestingly, in this 1992 book, wherein Williams spends a good deal of time emphasizing the important causal roles of interactors in selection processes, he still prefers to call replicators (or 'codex'), and not interactors, the 'unit of selection', because they are the beneficiaries of the long-term selection process (1992, 16; cf. Lloyd 2001 on the beneficiary question).

5. Claims of Equal Representation. The fundamental claim of the pluralists is that anything that a hierarchical selection model can do, a genic selection model can do just as well. Thus, much attention is paid to showing that the two types of models can represent certain patterns of selection equally well, especially those that are conventionally considered hierarchical selection exemplars.

5.1. Specific Claims of Equivalence. Let's take a closer look at S&K's example, the sickle-cell anemia case, in which the heterozygote is superior to either homozygote in malarial environments. This is usually described as the heterozygotes having a higher fitness than the other pairs, and thus that selection is occurring on the level of the genotype.

Now look at the alleles in the heterozygote superiority case from the genic point of view. "The alleles form parts of one another's environments. . . . The property of directing the formation of a particular kind of hemoglobin, has a unique environment-dependent effect on survival and reproduction" (S&K 1988, 345) (see Waters 1991, 560). The key lies in considering the other allele at a locus as a crucial part of the focal allele's environment, and calculating allelic fitnesses according to these crucial environment parts.

S&K work it out this way in the sickle-cell case:

Let P1 be the collection of all those allele copies that occur next to an *S* allele, and let P2 consist of all those allele copies that occur next to an *A* allele. Then, the property of being the *A* gene (i.e., property of directing the production of normal hemoglobin) has a positive effect on the production of copies in the next generation in P1, and conversely in P2. "In this way, we are able to *partition* the population and to achieve a Dawkinsian redescription" (1988, 347). Thus, an allelic-level description is derived, once the population is divided into the right subenvironments—ones in which the alleles have noncontext-sensitive fitness effects.

Note that S&K's redescription requires the values for the positive effects of being an *A* gene in population P1, and the negative effects of being an *A* gene in population P2. These numbers are commonly known as the genotype fitnesses. On S&K's story, the fitness of allele *A* is $W(A) = pw(AA) + qw(AS)$. Thus, they need to have fitness information concerning what is usually considered the (higher-level) interactor in this system, the organism or genotype (*AS*), in order to derive their Dawkinsian redescription. The reason this is important is because they have offered no principled way of telling when the higher-level causal fitness parameters will be important to their allelic model.

In a paper that some take to support the S&K genic claims, Peter Godfrey-Smith and Richard Lewontin (1993) proved that they could create regular population genetics models based on either allelic frequencies

or genotypic frequencies. They formalize some of the results implicit in S&K's suggestion for remodeling. The results are quite interesting, because in Lewontin's ingenious derivation of an allelic from a genotypic model, we find, sitting right in the midst of the 'allelic' model, a genotypic fitness parameter. This result (inadvertently) provides further support for the view that empirically adequate allelic-level models are dependent *in their construction* on higher-level information. This is a great deal more serious than simply obtaining a parameter value from another model, which happens frequently; rather, S&K have no model at all, without relying on the entire higher-level structure.

Waters acknowledges this issue. In his discussion of the Sober and Lewontin model (1982) of heterozygote superiority, he observes that "changes in gene frequencies are determined by the fitnesses of gene pairs . . . hence, it appears that the force of selection impinges on gene complexes, not on individual alleles" (1991, 557). Waters notes that the allelic fitness of gene *S* with *A* in its environment may have the same *numerical value* as the fitness of diploid genotype *AS*, "but the *interpretations* are not the same. . . . One concerns the propensity of a single gene to make good in a genetic environment, the other concerns the propensity of a gene pair in a less inclusive environment" (1991, 560). But what does it signify? If the parameters are semantically distinct and you must use the higher-level information, then the pluralists' models are parasitic, derivative, and hence, not independent. We still need the information about heterozygote fitnesses, and we need to get it the same way—by looking for interactors in a selection process. Thus, the S&K account is not a genuine alternative to the hierarchical account; it is simply a renaming of parts of the mathematical structure developed through hierarchical means, i.e., it is derivative. KSW may object that they are not obliged to furnish new methods or rules of thumb for determining the fitness of a gene. But remember, they are claiming to have a new genic-level theory. They have eschewed the units of selection debates, and presumably, their methods and procedures, having claimed that they are all 'muddled metaphysics' and 'pseudoproblems'. Thus, they cannot just appeal to ordinary interactor-locating devices or methods that arise squarely out of the very debates they have trashed.

Let us now turn to one of the classic cases of the efficacy of interdemic, or group, selection—the case that even Williams acknowledged was hierarchical selection.[18] Lewontin and Dunn, in investigating the house mouse, found first, that there was segregation distortion, in that well over

18. Note that this is not a case of trait group selection, and, as such, is not subject to any of the mathematical intertranslatability arguments recently highlighted in the philosophical literature.

80% of the sperm from mice heterozygous for the *t*-allele also carried the *t*-allele, whereas the expected rate would be 50%. They also found that male homozygotes with two *t*-alleles were sterile. But there was a further complication, which is that, even taking into account the level of the *t*-allele segregation distortion, plus the fact that there was strong selection against *t*-allele homozygotes, *t*-alleles tended to occur at a lower frequency in populations of house mouse than was expected. Consequently, given that they knew the *biology* of the house mouse tended to favor small breeding groups over large ones, Lewontin and Dunn investigated the effects of differential group extinction. They found a substantial effect of group extinction based on the fact that female mice would often find themselves in groups in which all the males were homozygous for the *t*-allele, and hence sterile, and the group itself would therefore go extinct. Thus, they developed a hierarchical selection model on which three levels of interactors were operating simultaneously. This, then, is how a genuine, empirically robust, hierarchical model was developed (Lewontin and Dunn 1960; Lewontin 1962).

What the genic pluralists want to note about this case is very narrow, that is the question "whether there are real examples of processes that can be modeled as group selection can be asked and answered entirely *within the genic* point of view" (KSW 1990, 160; emphasis added). Waters tells how to "construct" a genic model of the causes responsible for the frequency of the *t*-allele (1991, 563).

In order to determine the fitness parameter of a specific allele, let's call it *A*, we would need to know what kind of environment it is in at the allelic level, e.g., if it is paired with a *t*-allele. Then we would need information about a further distinct detailed layer of the environment of *A*, such as what the sex is of the 'environment' it is in. If it is in a *t*-allele arrangement, and it is also in a male environment, the allelic fitness of *A* would be changed, as a result of segregation distortion. And so on, with the demic environments, too. As we can see, various aspects of the allele's environment are built up from the gene out, depending on what would finally make a difference to the gene's fitness. The overall fitness of the *A* allele is calculated by adding up the fitnesses in each set of specialized, detailed environments and weighting them according to the frequency of each environment. Question: How does Waters know that interactions at the group or organismic level will have an effect on genic fitness?

Significantly, each one of these levels of genic environment is an interactor on the ordinary hierarchical view. This is how all the causal information from the regular hierarchical model gets transformed and derived into genic terms. So the end result is that what was represented as a causally relevant interactor in the hierarchical model ends up being renamed as a causally relevant allelic subenvironment type in the allelic

model. *Different state space, same overall fitness structures, same causes*. (See Section 6, Premise 2.)

Waters insists that the empirical issues do not disappear under his genic analysis. For instance, the possibility that female mice could be caught in populations in which all males are homozygous for the *t*-allele "is as much an issue for the genic selectionist as it is for the group selectionist" (1991, 564). Thus, Waters writes, "I can see no basis for concluding that [this genic representation] misrepresents the causal process" (1991, 564). "What appears as a *multiple-level selection process* (e.g., selection of the *t*-allele) to those who draw the conceptual divide [between environments] at the traditional level, appears to genic selectionists of Williams's style as *several selection processes* being carried out at the same level within different genetic environments" (1991, 571; emphasis added). Note Waters' identification of selection processes with the renaming of parts of the hierarchical mathematical structure (see Section 6).

5.2. How Is It Done? This whole procedure of determining which level of allelic environment needs to be included looks suspiciously like those used to determine whether something is functioning as a hierarchical interactor. Given that the pluralist genic selectionists have eschewed the interaction question, its presence in the middle of their model is surprising. But, they need to know if there are aspects (traits) of the environment functioning at certain levels that make a difference to replicator success. How is that discovered? What are the heuristics for determining relevant partitioning of the environments?

The similarity between the hierarchical interactor and genic environmental ways of seeking information necessary for an adequate model is clinched by the fact that the pluralists want to use the *same tools* for delineating genetically relevant environments as others do when they are looking for interactors. The point here is not that there is something wrong with the genic selectionists wanting to use efficient tools for dividing up the allelic environments. Rather, the point is that they claimed to have overcome the 'quibbles' involving *just those issues*. Waters, however, suggests that a genic analysis could be based on my additivity approach to identifying *interactors*. The additivity criterion presupposes some definition of the environment. Change the way that the environment is defined, for example to Williams' way, suggests Waters, and "genic selectionists could individuate genetic environments" under the additivity approach" (1991, 563).[19] Unfortunately, the sensible notion of borrowing a method

19. My (1986, [1988] 1994) 'additivity criterion' for an interactor has been criticized by Peter Godfrey-Smith (1992) and Sahotra Sarkar (1994). But both Godfrey-Smith and Sarkar misread the additivity criterion as requiring that we have a level of selection

for identifying potential higher-level interactors in order to determine the genic environments, and thus to have more adequate genic-level models, embroils Waters in the interactor debates that he claimed to avoid using the genic approach.

In a more oblique move, S&K appeal to a traditional approach for identifying interactors in order to divide up genic environments for allelic models in an empirically adequate fashion. Brandon (1982) used the statistical idea of screening off to identify which levels of entities are causally effective in the selection process. In other words, it is a method used to isolate interactors using traditional notions of environments.[20] S&K, however, propose that screening off be used for the genic approach by changing the notion of environment to the allelic environment (1988, 354). So, despite the fact that S&K, like Waters, claim to overcome the units of selection 'pseudoproblems', they, like Waters, end up taking sides in the interactor debate.

For Waters, and more obliquely S&K, levels of interaction important to the outcome of the selection process (in genic terms) are being discovered in the usual ways—that is, by using hierarchical approaches to identify various levels of interactors, and that information is then being translated into talk of the differentiated and layered environments of the genes.

Given this derivative method of model building, the crucial question for KSW is whether the problem of interactors has really been disposed of. The genic view requires a hierarchical set of environments in order to develop a workable genic fitness parameter. Thus, anything that makes a difference to genic fitness must be partitioned off into a separate environment of the gene. This is done in their examples in only one way—by taking the already established causally based higher-level model structures, and by terminological transformation and mathematical derivation, 'converting' them into genic environments (Godfrey-Smith and Lewontin 1993; S&K 1988; Waters 1991).

Note, in particular, that the relevant *causes* remain at the hierarchical level, e.g., Waters' inclusion of demic 'environments' as causally relevant. In other words, it is the demes' properties interacting with the demes' environments that are being included here as causally relevant to allelic success; these are read straight off of the hierarchical structures' inter-

---

not only if there is a nonadditive component of fitness at that level, but also if all parameters in the model, including dominance and even the genic-level fitness parameter, all have nonzero additivity. This is not what the definition says, as is clear from reading the rest of Chapter 5 ([1988] 1994). However, I am not currently defending any specific approach to the interactor question. For an insightful account of my additivity criterion, see Griesemer and Wade (1988).

20. Brandon (1990) distinguishes between physical, ecological, and selective environments.

pretations of causes. But, what has happened here besides a derivative change in state space and the altering of names from 'interactor' to 'environment part'? Standard hierarchical interactor definitions and techniques are used to isolate, divide, and structure the 'environments' causally relevant to genic success, and renaming these higher-level investigations and causes does not make them go away. This is especially important to see in the case of causes; when hierarchical structures' causes are used to produce empirically adequate lower-level models, the fact that the higher-level causes are now 'hidden' does not mean they play no role in the new models. More importantly, because the lower-level models are fully derivative from the hierarchical models, there are no *new* causes introduced; there are only causes that are derivative from the hierarchical structures.

The reader may be wondering, at this point, whether the pluralists' commitment to using interactor methods and models makes any difference to their broader claims, which are ontological. Their exact relations to the alternative account, the hierarchical account, then become the center of attention.

Two models that are mathematically equivalent may be semantically different, that is, they have different interpretations. Such models can be independent from one another, or be one derivative from the other. In the genic selection case, the pluralists appear to be claiming that the genic-level models are independent from the hierarchical models. The claim is that although the genic models are mathematically equivalent, they have different parameters, and a different interpretation, and they are completely independent from hierarchical models.

But, despite the pluralists' repeated claims, we can see *from their own calculations and examples* that theirs are *derivative* models, and thus, that their 'genic' level causes are derivative from and dependent on higher-level causes. Their genic-level models depend for their empirical, causal, and explanatory adequacy on entire mathematical structures taken from the hierarchical models and refashioned. This is why their implicit or explicit use of the hierarchical interactor definitions is so damaging to their case that genic models are somehow competitors to hierarchical models.

This point becomes transparent when pluralist genic selectionists use techniques for locating hierarchical interactors in order to make their genic selection models work. By taking a stand on which specific technique to use in determining the levels of causal interaction (or equivalently, allelic environment) having significant effects on the genic level, the pluralist genic selectionists have (inadvertently) taken a stand regarding the very squabbles that they claim to avoid. And, they have taken sides in them without undertaking the considerable theoretical work involved in justi-

fying the choices they have made. Thus, they occupy positions that they are unable to defend, due to their denial that there is a problem at all.

Given the failure of the pluralist selectionists to extricate themselves from the interactor debates, and their cavalier and undefended adoption of one or another hierarchical definition of an interactor in order to make their genic models empirically adequate, the punch line to their paper seems especially ill-conceived. They argued that the presence of equally adequate representations of evolutionary processes means that philosophers and biologists can cease 'quibbling' about units of selection. Instead, it seems: After the genic selectionists take an undefended position on the units of selection 'quibbles', we must allow the possibility of many, equally adequate representations of evolutionary processes, including genic ones.

6. The General Pluralist Argument. So far, I have focused on how the genic pluralists' models are derivative from higher-level models, in particular their dependence upon prior solutions to the units of selection problem (in this case, the interactor question). This might seem to leave their general arguments about the existence of alternate models, and the seemingly arbitrary nature of choice among them, untouched. But, this is not so. The issues of derivativeness, dependence, and the general pluralist argument are intimately connected. The basic structure of the argument for pluralism, along with evaluations of its premises, are set out below.

> **Premise 1**. "There are alternative, *maximally adequate*, representations of the causal structure of the selection process" (S&K 1988, 358; emphasis added). (Cf. Waters 1991, 572.)

Analysis. Derivativeness implies that these alternative representations are not genuine alternatives. Rather, they are semantic reinterpretations dependent on the empirical and explanatory adequacy of higher-level models. Investigators into higher-level interactors incorporate environmental factors, trait-bearing entities, and demonstrations of correlations with fitness components, etc., (as reviewed in Section 4) into their models. This results in mathematical structures, (however represented) key aspects of which are required for the empirical adequacy of the structures. Renaming parts of these structures does not change conclusions drawn about other parts of these structures, particularly the level(s) at which entities interact with their environments through their traits (traditionally, the definition of an interactor or unit of selection). Premise 1, therefore, is undermined by the derivative nature of the multienvironmental allelic structures, and especially by the derivative nature of any causal claims. There are no alternative causes, nor operative levels of selection; there are simply renamed structures in derived models.

Subpremise 1. "Specifically, we can always find a way to present a selection process in terms of the *causal efficacy* of genes" (S&K 1988, 358; emphasis added).

Analysis. This premise, a special case of Premise 1, is undermined by Lewontin's argument, which shows the dynamical insufficiency of presenting selection processes purely in terms of genic causes in genic state space and reconstructed allelic environments. The premise is already undermined by the derivativeness discussed in the analysis of Premise 1. Lewontin gives an in-principle argument against the possibility of such allelic models. (See Section 8.)

Premise 2. (Partly suppressed.) Genic and hierarchical accounts identify different causal efficacies or agents or causes (S&K 1988, 358; Waters 1991, 562).

Analysis. This premise appears to result from metaphysical assumptions involving agents. Specifically, once a model is reformulated in terms of a particular state-space in which alleles are the state variables, there seems to be a tendency to identify a distinct causal agent. An example of this metaphysical move occurs in Waters' account of sickle-cell anemia, wherein the 'agent' in the model to which we attribute 'causal power', as well as a different 'level of selection', is the allele, because it is the state space of the model (Waters 1991, 564). Waters gives no defense of this position. Williams defends it by appealing to what I have analyzed as the 'beneficiary question': What entity is around in the evolutionary long run to *benefit* from an evolution by selection process? In sexually reproducing species, the answer must be the allele; hence, according to Williams (1966), the allele is the 'unit of selection'.

S&K, in contrast, do give arguments for the differences in causal attribution of the hierarchical and genic-level models, but the arguments they appeal to are Dawkins', nearly all of which are aimed at organismic selectionists, which are not under consideration here. (See Lloyd [1988] 1994, 2001 for discussion.) Dawkins' arguments apply to the replicator question and not to the interactor debate, and hence are misapplied by S&K.[21] Dawkins' own argument is also based fundamentally on his interest in the 'units of selection' as beneficiaries, following the lines of Williams, and fares no better against those who are engaged in the interactor debates (Lloyd 2001). The problem with focusing on the 'units of selection' as beneficiaries question in the late 1980s is that all of the

21. Sterelny and Kitcher repeat Dawkins' arguments that the focus should be on the 'causal properties' of alleles, but their argument is undone by their definition of allelic fitness, which falls to Lewontin's objection (S&K 1988, 346).

debate in population genetics, and the vast majority of the debate in philosophy, focused clearly on either the 'interactor' question, or a combination of the 'interactor' question with the 'manifestor of adaptation' question. Thus, KSW, in addressing the 'unit of selection quibbles' relating to population genetics, can only plausibly be read as addressing the genetics arguments and related philosophical arguments concerning specific disagreements about how interactors or interactors-and-manifestors-of-adaptations are defined and identified. Any appeal to the beneficiary question would be, at best, off-point—a metaphysical sideshow.

But perhaps we do not have to appeal to metaphysics at all. There is a precedent in scientific practice for changing causal accounts when the state space is changed. As Rob Cummins (personal communication, 2005) has noted:

> If you have discriminations that you can make empirically, yet there doesn't seem to be room in your current state space to allow them to do some work, it is a standard practice in the quantitative sciences to move up to a bigger state space, simultaneously changing your interpretation of the model's causal structure. The same thing can work in the downward direction. Thus, we have a principle—*change state space: change causal structure*.

These cases, however, are purely empirically driven.

With the genic pluralists, in contrast, we change the state space, but not the empirical content of the theory at all, thus undermining the justification for accompanying a change in state space by a change in causal structure. The principle stated above holds where there are changes in empirical content.[22] Thus, the pluralists are trying to have it both ways. An unwary reader, however, may easily be misled by this particular move made by KSW.

Thus, we are back to the metaphysical claims discussed above. In sum, much of the genic pluralists' view rests on implicit, heavily loaded metaphysical assumptions in which causal agency is assigned according to the state space in the model: If you have a model in which the allele is the entity with a fitness parameter, then assume that the allele is the 'actor', 'agent', or 'cause' of the selection scenario.

> **Conclusion**. Genic pluralism is correct (S&K 1988, 358). "[T]he causes of one and the same selection process can be correctly described by

22. Based on a discussion with Rob Cummins, January 27, 2005.

accounts which model selection at different levels" (Waters 1991, 572).[23]

Corollaries of Conclusion. (*a*) There is no units of selection problem; it has been dissolved by the availability of multiple models. (*b*) There are no targets of selection.

Analysis. The corollaries of the conclusion have been shown to be incorrect (see Sections 4 and 5).

Conclusion. Since Premise 1, Subpremise 1, and Premise 2 are undermined, the conclusion is unsupported. In addition, two corollaries of the conclusion have been shown to be false, which directly indicates the falsity of the conclusion.

7. Pragmatic Virtues. In presenting their claims for the importance of a completely genic-level evolutionary theory, S&K contrast 'two images' of evolution by selection. The two images they have in mind are organismic and genic selectionism. They state that the relative worth of the two images turns on two theoretical claims in evolutionary biology:

1. Candidate units of selection must have systematic causal consequences. If *X*s are selected for, then *X* must have a systematic effect on its expected representation in future generations.
2. Dawkins' genic selectionism offers 'a more general theory of evolution', one that can handle organismic selection cases, but also other cases that are problems for organismic selection. (1988, 340)

Challenges to pluralist genic selectionist accounts typically assert that pluralism fails to meet Condition 1. Pluralists' arguments purportedly showing the equivalence of genic and higher-level models are aimed at rejecting this critique. As I have shown, this defense comes at a very high price, namely, an explicit appeal (via adoption of hierarchical methods and models) to higher-level causal consequences that the genic pluralists claim to avoid.

Additionally, S&K's arguments for the superior generality of the genic approach, aim at the wrong target, organismic selection, since none of

23. This is a legitimate conclusion in the small number of selection cases found that can be categorized as 'extended phenotype' cases. But these cases make up a quite small minority of selection cases, and thus cannot be used to represent hierarchical selection as a whole. Dawkins' mistake, in his 1982 book, was to take only organismic selection and genic selection as his two alternatives; thus, he ended up concluding that there was an abundance of extended phenotype cases. Once hierarchical selection is taken into account, this clearly doesn't hold.

the authors they attack actually defend that view, as Sober pointed out (1990, 152). When S&K do finally confront the real contender, hierarchical selection theory, they seem to misunderstand it (1988, 359). Their mistakes here are identical to their mistakes in the arguments for the other putative virtue of the genic approach, its unity, so we will skip ahead to that virtue, since at least it's aimed at an appropriate target.

Genic selectionism's other ostensible virtue is that it is more unified than the hierarchical view, according to Dawkins and KSW. With genes as both replicators and interactors, the entire theory of evolution by natural selection is unified and all selection phenomena can be taken into account. Against this they compare the hierarchical account, where "processes are diverse in the kinds of representation they demand" (1988, 354). Thus, it appears that you have need for one type of representation for group selection, another for kin selection, another for genic selection, and so on. Hierarchical selection looks like a big mess. But this reveals a deep misunderstanding of hierarchical theory, which was unified in its initial form (ironically) by Dawkins, and then in its current form by Hull.

According to Hull, at each level of the biological hierarchy where there is selection going on, there is an interactor at that level and a replicator at that level, or more usually at a lower level. The biology of the situation determines which entities play the functional roles of interactors and replicators in a given case (1980). The hierarchical theory is thus a highly unified theory, capable of accounting for all cases of selection, including genic selection.[24] The formal models for hierarchical selection illustrate this unification beautifully.[25]

8. Conclusion. In sum, the genic view does not have the big advantage claimed; it does not avoid the unit of selection 'quibbles' after all. But, the causal questions concerning which entities are interacting with their environments in a way that affects replicator success—including questions concerning how to know when to take such interactors into account and how to detect them—are crucial for the genic approach, as the genic pluralists themselves have shown. The causal equivalence of the models is better thought of as derivative; higher-level causes are reformulated down in model translation and imported into the genic models, not changing one whit the level at which the identified interaction between trait and environment that makes the difference to genic success (the interactor)

24. Which, by the way, Dawkins couldn't do with his replicator-and-vehicle theory. See footnote 10.

25. E.g., Arnold and Fristrup 1982; Heisler and Damuth 1987; Damuth and Heisler 1988; Wade 1985.

actually appears in the overall mathematical structure from which the genic model is derived.

This means that the crucial question regarding the levels of interaction necessary to model a given selection process is still outstanding, whether it is put in hierarchical terms or genic ones. Worse, the pluralists we examined attempt to sidestep this question, but cannot; they end up taking sides inadvertently in the interactor debate, all the while claiming they have overcome just that debate. Neither philosophers nor biologists can ignore the debate about interactors in selection; the outcomes make a difference to the empirical, explanatory, and causal success of the models produced.

Pluralist genic selectionists also claim that once genic and other alternative representations of selection processes are available, there is no scientific way (except perhaps the reductionist one) to choose among the various alternative models (S&K 1988, 358).[26] In addition to the derivativeness argument, the additional flaw with this argument is that Lewontin has shown the dynamical insufficiency of their approach to genic models, wherein the allelic frequencies and the layered allelic environments of the alleles determine fitnesses. Even if they had an independent causal definition at the genic level (which they do not), the models would not be the promised "[presentations of] a selection process in terms of the causal efficacy of genes" (S&K 1988, 358). Thus, there can be no 'causal equivalences' of the sort they claim to have shown.

Let me spell this out in more detail. Perhaps S&K, especially, have simply chosen badly in illustrating their point, and perhaps the revamping of Williams that Waters is suggesting could do the trick, without relying on the derivation from hierarchical models. Not so, says Lewontin. There is something fatally wrong with any model that attempts to model a selection process using genic selection parameters and associated allelic environments. Indeed, there is also something faulty in using purely genotypic state space models to represent selection, since, at every generation, we need allelic-level information about segregation and usually, recombination, in order to produce a pool of gametes. Thus, properly conceived, population genetics models need, at a minimum, information from both the allelic and the genotypic space.

Since all regular population genetic models require both allelic and genotypic level models to produce dynamically workable models, Lewontin argues the following. Take an allelic-level model in which we are using allelic-environments to determine allelic-level frequencies (as all the models suggested by the pluralists require). At each generation, we must

26. Waters (1991) is probably an exception to the reductionist move, here.

readjust the selection coefficient according to the frequencies of the different alleles. How is this done? The obvious way is to take a weighted average of the fitnesses in the different allelic 'environments', weighted by how frequent the different 'environments' are (as suggested by S&K 1988, 346). But this will only work if the alleles find themselves in the different environments randomly, i.e., if we are in Hardy Weinberg proportions. So we are required to determine how frequent the combinations of alleles and 'environments' are, i.e., what the frequencies of different diploid genotypes are. That is, we are required to go up to the higher level; it can't all be done at the allelic level, any more than genotypic calculations can all be done at the genotypic level alone—they are reliant on the allelic level.

Even if the population is in Hardy Weinberg equilibrium, that must be confirmed, which also requires going up to the genotype level. So, no matter what, in order to calculate the next generation of allele frequencies, we must know the next generation's genotype frequencies, in order to calculate the fitnesses of the alleles in the allelic space. The only alternative is to stay in the allelic space completely, and track empirically the changes in the genic selection coefficients at each generation, in which case, every generation would produce a new model, unlinked to the model of the previous generation. If we were to approach it this way, there would be no science of population genetics, i.e., no way to produce trajectories of gene changes over time, or predictions of changes, or explanations of changes. Consideration of more than one locus for selection (as many of the models touted by the genic selectionists involve) makes the empirical tracking approach mathematically intractable, as well as predictively useless. Thus, *there is no way to actually represent a selection process over time with an allelic-level model*. They are not dynamically adequate without information from at least the genotypic level, and maybe higher levels, so even setting aside the issue of the origins of the allelic models from hierarchical or genotypic models, the allelic models cannot be used at all without the constant input of genotypic or hierarchical models in order to represent selection processes. There is no equivalence here, only utter dependence.[27]

Obviously, this argument completely destroys Subpremise 1. But even if Lewontin's argument were to fail, and through some other form of modeling, an independent purely genic environmental causal model could be rendered empirically adequate, KSW have certainly not established this. Meanwhile, there are no 'causal equivalences', and so no independent genic models. Thus, there are no independent, maximally adequate, caus-

27. Based on personal correspondence with Richard Lewontin, January 15, 2005.

ally different accounts of any particular selection episode to choose between: There are no independent genic causal models, and all the other models offered as 'alternative views' are derivative. When combined with the problems with Premise 2, the genic pluralists are left in trouble.

Not only have the genic pluralists failed to show that the units of selection question is obsolete, they have failed to show their famous 'equivalence' between alternate causal models of selection, particularly the genic ones: Generally, such supposed 'alternative causal models of selection' are (1) mere derivations of hierarchically structured models and not causal alternatives as advertised, or (2) as Lewontin shows, nonderivative, but dynamically inadequate.

The only sustainable conclusion of 'The Return of the Gene' and 'Tempered Realism' turns out to be that accurate, causal, empirically adequate genetics models *are precisely* the hierarchical structures that KSW have relabeled.

Thus, rather than the very weak (and unworkable) equivalence form of pluralism that KSW find in the units debates (and in the genic reductionism of S&K), I see four factors pointing toward of one of the stronger forms of pluralism mentioned in Section 2: the *t*-allele case; the methods for isolating and identifying interactors; proven derivativeness; and Lewontin's argument for the unsustainability of an allelic-level model-type. In particular, descriptions of some or all of the different levels of a system are required for an adequate scientific account (see Section 2). This hierarchical form of pluralism instantiates locally, for a given phenomenon, a nonreductive form of unity, because all the different forms of descriptions are brought to bear. Thus, hierarchical selection exemplifies a strong and important form of pluralism, exactly contrary to the claims made by KSW; far from being a form of 'monism', as charged, it is a strong form of pluralism, standing in contrast to the weakened (and reductionist, in the hands of S&K) form of pluralism advocated by the genic pluralists.

Sometimes simpler models are useful as tools, as has long been recognized by all geneticists. (For recent discussion, see Michod 1999.) But playing with the genetics in such a way that it doesn't improve the empirical adequacy of the models or theoretical understanding of evolutionary change is simply a step backwards to 1966.[28] Remember, Williams'

28. Kerr and Godfrey-Smith's (2002) article claims to do more than this; specifically, they believe viewing an evolutionary process from both individual and trait group points of view yields insight that might not otherwise be available. Note that their paper helps itself to the higher-level information from the beginning, clearly begging an important empirical and methodological question. Nevertheless, it is also heavily indebted to Eshel's work from the 1970s—work that lies firmly in the tradition discussed in the present paper, and so might well be compatible with many of the conclusions drawn here.

(1966) allelic parsimony argument was widely interpreted as a reason to never even take a look at higher-level interactions, interactions that eventually led to empirically more adequate models (Lloyd [1988] 1994, 94–96; Williams 1966, 5, 66, 93–95; Williams 1992). Similarly, the genetic pluralists seem to want to return us to a pre-1980s state of population genetics, one in which structured population genetics and hierarchical selection structures played little or no role, empirically or theoretically.

In the end, confused metaphysics (or twisted scientific practice, take your pick) are insufficient to fill the gaps in the arguments the genic pluralists present. Thus, genic pluralism fails on multiple counts; it serves only to highlight the success of hierarchical approaches to selection processes.

REFERENCES

Arnold, Anthony J., and Kurt Fristrup (1982), "The Theory of Evolution by Natural Selection: A Hierarchical Expansion", *Paleobiology* 8: 113–129.

Arnold, Stevan J., and Michael J. Wade (1984), "On the Measurement of Natural and Sexual Selection: Applications", *Evolution* 38: 720–734.

Brandon, Robert N. (1982), "The Levels of Selection", in Peter D. Asquith and Thomas Nickles (eds.), *PSA 1982: Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, vol. 1. East Lansing, MI: Philosophy of Science Association, 315–323.

——— (1990), *Adaptation and Environment*. Princeton, NJ: Princeton University Press.

Cat, Jordi (1998), "The Physicists' Debates on Unification in Physics at the End of the 20th Century", *Historical Studies in the Physical and Biological Sciences* 28: 253–299.

——— (2000), "Must Microcausality Condition Be Interpreted Causally? Beyond Reduction and Matters of Fact", *Theoria* 37: 59–85.

——— (2004). *Physics beyond Laws and Theories: The Limits of Unity, Universality and Precision*. Forthcoming.

Damuth, John, and I. Lorraine Heisler (1988), "Alternative Formulations of Multilevel Selection", *Biology and Philosophy* 3: 407–430.

Darden, Lindley, and Nancy Maull (1977), "Interfield Theories", *Philosophy of Science* 44: 43–64.

Dawkins, Richard (1982), *The Extended Phenotype*. New York: Oxford University Press.

Dugatkin, Lee A., and H. Kern Reeve (1994), "Behavioral Ecology and Levels of Selection: Dissolving the Group Selection Controversy", *Advances in the Study of Behavior* 23: 101–133.

Dupré, John (1993), *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Glennan, Stuart (2002), "Contextual Unanimity and the Units of Selection Problem", *Philosophy of Science* 69: 118–137.

Glymour, Bruce (1999), "Population Level Causation and a Unified Theory of Natural Selection", *Biology and Philosophy* 14: 521–536.

Godfrey-Smith, Peter (1992), "Additivity and the Units of Selection", in David Hull, Micky Forbes, and Kathleen Okruhlik (eds.), *PSA 1992: Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*, vol. 1. East Lansing, MI: Philosophy of Science Association, 315–328.

Godfrey-Smith, Peter, and Richard Lewontin (1993), "The Dimensions of Selection", *Philosophy of Science* 60: 373–395.

Goodnight, Charles J., and Lori Stevens (1997), "Experimental Studies of Group Selection:

What Do They Tell Us about Group Selection in Nature?", *American Naturalist* 150: S59–S79.
Gould, Stephen J. (2002), *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.
Griesemer, James (2000), "Development, Culture, and the Units of Inheritance", *Philosophy of Science* 67 (Proceedings): S348–S368.
Griesemer, James, and Michael J. Wade (1988), "Laboratory Models, Causal Explanations and Group Selection", *Biology and Philosophy* 3: 67–96.
Heisler, I. Lorraine, and John Damuth (1987), "A Method for Analyzing Selection in Hierarchically Structured Populations", *American Naturalist* 130: 582–602.
Hull, David L. (1980), "Individuality and Selection", *Annual Review of Ecology and Systematics* 11: 311–332.
——— (1988), "Interactors versus Vehicles", in H. C. Plotkin (ed.), *The Role of Behavior in Evolution*. Cambridge, MA: MIT Press, 19–50.
Kawata, Masakado (1987), "Units and Passages: A View for Evolutionary Biology", *Biology and Philosophy* 2: 415–434.
Kerr, Benjamin, and Peter Godfrey-Smith (2002), "Individualist and Multilevel Perspectives on Selection in Structured Populations", *Biology and Philosophy* 17: 477–517.
Kitcher, Philip, Kim Sterelny, and C. Kenneth Waters (1990), "The Illusory Riches of Sober's Monism", *Journal of Philosophy* 87: 158–161.
Lande, Russell, and Stevan J. Arnold (1982), "The Measurement of Selection on Correlated Characters", *Evolution* 37: 1210–1226.
Lewontin, Richard (1962), "Interdeme Selection Controlling a Polymorphism in the House Mouse", *American Naturalist* 887: 65–78.
——— (1970), "The Units of Selection", *Annual Review of Ecology and Systematics* 1: 1–18.
Lewontin, Richard, and Leslie Dunn (1960), "The Evolutionary Dynamics of a Polymorphism in the House Mouse", *Genetics* 45: 705–722.
Lloyd, Elisabeth A. (1986), "Evaluation of Evidence in Group Selection Debates", in Arthur Fine and Peter Machamer (eds.), *PSA 1986: Proceedings of the 1986 Biennial Meeting of the Philosophy of Science Association*, vol. 1. East Lansing, MI: Philosophy of Science Association, 483–493.
——— (1992), "Unit of Selection", in Evelyn F. Keller and Elisabeth A. Lloyd (eds.), *Keywords in Evolutionary Biology*. Cambridge, MA: Harvard University Press, 334–340.
——— ([1988] 1994), *The Structure and Confirmation of Evolutionary Theory*. Reprint. Originally published by Greenwood Press. Princeton, NJ: Princeton University Press.
——— (2001), "Different Questions: Levels and Units of Selection", in Rama S. Singh, Costas B. Krimbas, Diane B. Paul, and John Beatty (eds.), *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*. New York: Cambridge University Press, 267–291.
Maynard Smith, John (1987), "How to Model Evolution", in John Dupré (ed.), *The Latest on the Best: Essays on Evolution and Optimality*. Cambridge, MA: MIT Press, 119–132.
Mayo, Deborah, and Norman Gilinsky (1987), "Models of Group Selection", *Philosophy of Science* 54: 515–538.
Michod, Richard (1999), *Darwinian Dynamics*. Princeton, NJ: Princeton University Press.
Nunney, Len (1985), "Group Selection, Altruism, and Structured-Deme Models", *American Naturalist* 126: 212–230.
Okasha, Samir (2004), "Multilevel Selection and the Partitioning of Covariance: A Comparison of Three Approaches", *Evolution* 58: 486–494.
Price, George R. (1972), "Extension of Covariance Selection Mathematics", *Annals of Human Genetics* 35: 485–490.
Sarkar, Sahotra (1994), "The Selection of Alleles and the Additivity of Variance", in David Hull, Micky Forbes, and Richard M. Burian (eds.), *PSA 1994: Proceedings of the 1994 Biennial Meeting of the Philosophy of Science Association*, vol. 1. East Lansing, MI: Philosophy of Science Association, 3–12.

Shanahan, Timothy (1997), "Pluralism, Antirealism, and the Units of Selection", *Acta Biotheoretica* 45: 117–126.
Sober, Elliot (1984), *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, MA: MIT Press.
——— (1990), "The Poverty of Pluralism: A Reply to Sterelny and Kitcher", *Journal of Philosophy* 87: 151–158.
Sober, Elliot, and Richard Lewontin (1982), "Artifact, Cause and Genic Selection", *Philosophy of Science* 47: 157–180.
Sober, Elliot, and David S. Wilson (1994), "A Critical Review of Philosophical Work on the Units of Selection Problem", *Philosophy of Science* 61: 534–555.
——— (1998), *Unto Others*. Cambridge, MA: Harvard University Press.
Stanford, P. Kyle (2001), "The Units of Selection and the Causal Structure of the World", *Erkenntnis* 54: 215–233.
Sterelny, Kim, and Philip Kitcher (1988), "The Return of the Gene", *Journal of Philosophy* 85: 339–361.
Van der Steen, Wim J., and Hugo A. Van den Berg (1999), "Dissolving Disputes over Genic Selectionism", *Journal of Evolutionary Biology* 12: 184–187.
Wade, Michael J. (1978), "A Critical Review of the Models of Group Selection", *Quarterly Review of Biology* 53: 101–114.
——— (1985), "Soft Selection, Hard Selection, Kin Selection, and Group Selection", *American Naturalist* 125: 61–73.
Waters, C. Kenneth (1986), *Models of Natural Selection: From Darwin to Dawkins*. Ph.D. Dissertation. Bloomington, IN: Indiana University.
——— (1991), "Tempered Realism about the Force of Selection", *Philosophy of Science* 58: 553–573.
Williams, George C. (1966), *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.
——— (1992), *Natural Selection: Domains, Levels & Challenges*. Princeton, NJ: Princeton University Press.
Wilson, Robert A. (2003), "Pluralism, Entwinement, and the Levels of Selection", *Philosophy of Science* 70: 531–552.
Wimsatt, William (1980), "Reductionist Research Strategies and Their Biases in the Units of Selection Controversy", in T. Nickles (ed.), *Scientific Discovery: Case Studies*. Dordrecht: Reidel, 213–259.
——— (1981), "Units of Selection and the Structure of the Multilevel Genome", in Peter D. Asquith and Ronald N. Giere (eds.), *PSA 1980: Proceedings of the 1980 Biennial Meeting of the Philosophy of Science Association*, vol. 2. East Lansing, MI: Philosophy of Science Association, 122–183.