The Generational Cycle of State Spaces and Adequate Genetical Representation*

Elisabeth A. Lloyd, Richard C. Lewontin, and Marcus W. Feldman†

Most models of generational succession in sexually reproducing populations necessarily move back and forth between genic and genotypic spaces. We show that transitions between and within these spaces are usually hidden by unstated assumptions about processes in these spaces. We also examine a widely endorsed claim regarding the mathematical equivalence of kin-, group-, individual-, and allelic-selection models made by Lee Dugatkin and Kern Reeve. We show that the claimed mathematical equivalence of the models does not hold.

1. Introduction. A major struggle in the philosophy of evolutionary and population genetics has concerned the question of what the appropriate units should be for the dynamics of evolutionary change in sexually reproducing organisms. We propose an analysis of what needs to be known for an adequate genetical representation of evolutionary change. We then address a widely influential paper that claims that kin-, group-, individual-, and allelic-selection models are mathematically equivalent.

2. State Spaces. A useful tool for the analysis of the problem of evolutionary change is the state space. A state space is characterized by: (1) *entities* such as genes, genotypes, phenotypes, groups of individuals, and so on, and (2) *attributes* of these entities which may be either nonmetric

*Received January 2007; revised April 2008.

†To contact the authors, please write to: Elisabeth Lloyd, Department of History and Philosophy of Science, 130 Goodbody Hall, Indiana University, Bloomington, IN 47405; e-mail: ealloyd@indiana.edu; Richard Lewontin, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138; Marcus Feldman, Department of Biological Sciences, Stanford University, Stanford, CA 94305; e-mail: marc@charles.stanford.edu.

Philosophy of Science 75 (April 2008) pp. 140–156. 0031-8248/2008/7502-0002$10.00
Copyright 2008 by the Philosophy of Science Association. All rights reserved.

descriptors like molecular composition or metric descriptors for each of which there is a metric *dimension*, of the space, that are sufficient to localize each individual entity to a point in the space (Lewontin 1974; Lloyd 1994). Dimensions might also be morphological or behavioral, or, if the entities are collections such as populations, they may be allele or genotype frequencies, or mean phenotype, for example. The actual process of genetic transition between generations involves, in general, transformations in time in spaces of both genic and genotypic entities, as well as mappings from one space to another during the complete generational cycle. It is certainly true that particular special cases can be adequately represented as transitions in fewer spaces, but these cases involve assumptions, sometimes hidden, about certain parameters of the material processes that are being represented.

The simplest predictive structure for Mendelian inheritance involves a transition from diploid to haploid and then back to diploid spaces. The development of genetics as a science has been possible only by the deliberate modeling of inheritance as a sequential transition between the diploid and haploid spaces during the process of heredity. This science was able to develop as a system of regular prediction only because Mendel's insight into the phenomenon of segregation was joined to the identification of chromosomes as physical linear arrays of factors which recombine with each other during meiosis by a simple exchange of these linear arrays to produce haploid gametes. This chromosomal linearity, maintained in the recombination process, and sorted as alternative arrays into haploid gametes, made it possible to create linkage maps which could predict the outcome of crosses between multiple genetic variants that had not previously been combined in the same individuals. Standard formal transmission genetics is built on an alternating sequence of events that move from a diploid genotypic space of description of the parents to a haploid genic space of description of the formation of gametes, and back to the diploid genotypic space of description of the next generation.

For some purposes it seems possible to choose arbitrarily between genic and genotypic spaces, and these have received a great deal of attention in the conflict over the appropriate level of description for changes in populations under natural selection. Changes in population composition are usually described as change in allele frequencies and for one gene locus it is well known that, using allele frequencies and the proportions of individuals of different genotypes, the entire change in population composition can be described as a change in allele frequencies resulting from different genic fitnesses, where those fitnesses are calculated as gene-frequency-weighted averages of fitnesses of the genotypes containing those alleles (Godfrey-Smith and Lewontin 1993; Waters 2005). Note that this definition of allelic fitness requires that the genotypic fitnesses have been

previously determined; thus the necessary information is imported into the allelic space model from genotypic space (Lloyd 2005).

2.1. The Spaces of Hardy-Weinberg. It is not generally realized that the Hardy-Weinberg equilibrium for one biallelic locus was derived both by Hardy and by Weinberg using the frequencies of the diploid genotypes, P, Q, and R, rather than the now familiar $p^2 : 2pq : q^2$ description in terms of the allele frequencies p and q, which was introduced by R. A. Fisher in 1918. But to derive the Hardy-Weinberg equilibrium in terms of p^2, $2pq$, and q^2, one begins with arbitrary genotypic frequencies and shows that after a generation of random mating of diploid genotypes followed by Mendelian segregation in the production of gametes, the resultant proportions of offspring genotypes are algebraically equivalent to p^2, $2pq$, and q^2. It is sometimes claimed that these proportions can be trivially derived by multiplying the gamete frequencies directly, making no use of the diploid genotypic space, but this depends on the hidden assertion that planktoid mating is identical in outcome with random mating of diploid genotypes followed by Mendelian segregation, a claim that has to be established by first mapping diploid mating frequencies into frequencies of haploid pairs at fertilization. Central to this mapping is the calculation of genic output from genotypic inputs, that is, the mapping of a diploid space into a haploid space. The steps in the derivation of the Hardy-Weinberg equilibrium for a biallelic locus in terms of allele frequencies and the demonstration that this occurs after a single generation of random mating are as follows:

1. In the original mating population the frequencies of the three genotypes are specified as D, H, and R;
2. The six possible types of mating ($D \times D$, $D \times H$, $H \times R$, etc.) are assumed to occur at random, so the frequencies of these matings are simply the products of the genotypic frequencies (D^2, $2DH$, $2HR$, etc.).
3. The offspring ratios from each of these matings are specified. It is still possible to stay in the diploid space by using Mendel's proportions without any commitment to an underlying segregation mechanism. Alternatively, the segregation mapping into the genic space in step (5) below can be carried out here, in which case the expressions for D', H', and R' in (4) below will be in terms of allele frequencies p and q. Whether or not the segregation mapping is carried out at this point, however, the, assumption of unbiased segregation in heterozygotes is fundamental, since if there is meiotic drive then Mendel's ratios are incorrect.

4. The sums of the output genotypes D', H', and R' are calculated over all matings. These turn out to be

$$D' = (D + H/2)^2 \; H' = 2(D + H/2)(R + H/2) \; R' = (R + H/2)^2.$$

5. To turn this genotypic description finally into an allele frequency description $p^2 : 2pq : q^2$, it must be established that $D + H/2$ and $R + H/2$ are the allele frequencies p and q. But this requires mapping genotype space into gene space by Mendel's Law of Segregation and this segregation must be unbiased. If it is not, then both parts of the Hardy-Weinberg law—that p remains unchanged from one generation to the next and that D, H, and R can be expressed as p^2, $2p$, and q^2 with $p = 1 - q = D + H/2$—are false.

2.2. The Generational Cycle. In addition to the above complications of Hardy-Weinberg, we must take into account the fact that genotypic and genic information alone are not determinative of the various forces operating to change population composition because many of these processes function on phenotypes rather than simply genotypes. Mating patterns, and probabilities of survival and reproduction, although influenced by genes are a consequence of developmental events that are contingent on the environment of the developing organism. In the most general case, environment includes influences of the phenotypes of previous generations by means of cytoplasmic inheritance through the egg. A complete general representation of genetic evolutionary processes then requires not two, but six spaces with sequential transitions within them and mappings from one to the other. (In what follows, *mappings* refer to translations or re-representations of a population or system in terms of a new set of state variables, i.e., in a new state space, while state *transitions* refer to dynamical laws that transform the system within a single state space from one step to the next.) These spaces are:

S_1 a diploid phenotypic space,
S_2 a diploid genotypic space,
S_3 a diploid pair phenotypic space,
S_4 a diploid pair genotypic space,
S_5 a haploid phenotypic space, and
S_6 a haploid genic space.

For the moment we consider only organisms that reproduce in discrete generations, with no overlap between generations. We can then begin to formalize the process of temporal change by choosing some point in the life cycle as the initial condition. For our purposes it is convenient to

choose the point of mating pair formation. The sequential transformations and mappings are then described in model stages as follows:

1. The description of the entities in phenotypic space (S_1) and their mapping into genotypic space (S_2): ($S_1 \rightarrow S_2$).
2. The formation of mating pairs by mapping pairs of entities in S_1 and S_2 into S_3 and S_4. The attributes of each individual, in the most general case, must include a list of ancestry relationships with other individuals so that any potential pair can be classified as sibs, cousins or any other descriptor of the ancestry tree. The mapping into mating pairs includes self-fertilizations or obligate outcrossings, and assortative mating by size, behavior, habitat chosen or degree of genetic relationship (inbreeding) or completely random mating in either dioecious or monoecious species, all of which can then be described genotypically ($S_1 \rightarrow S_3 \rightarrow S_4$ or $S_2 \rightarrow S_4$).
3. Mapping genotypes (diploid genomes) into genes (haploid sequences) by initial gamete formation within each individual, including genetic segregation, recombination, and mutation and meiotic drive, to form gamete potential subpools corresponding to each individual in S_4 ($S_4 \rightarrow S_6$).
4. Mapping potential gamete subpools into effective gamete subpools taking into account differential fertility of adult gamete producers, gamete motility and survival $S_6 \rightarrow S_5$.
5. Mapping into diploid genotypes taking into account differential compatibility of sperm with recipient ova $S_5 \rightarrow S_6 \rightarrow S_2$.
6. Mapping of diploid genotype into phenotype in development $S_2 \rightarrow S_1$.
7. Transition of the ensemble of phenotypes by survivorship to mating stage (S_1).
8. Start of a new generation in S_1, namely, S_1'.

While it is true that the system's transition from one generation to the next may be represented in any one of the seven model-stages and six spaces, we need to know the *entire loop*—and all the parameter values in each of the model stages in that loop—in order to obtain an accurate representation of the chosen state space in the next generation, that is, to get the state transformation equation between generations *within any one of the spaces* (e.g., if we want to move from (S_6) to (S_6')). For the transition in allelic space, we must move out of that space, into genotypic space to define the parameters, and back into allelic space in order to characterize the next generation. (We elaborate this point in 5.1 below.) The dynamical problems that arise in the context of several spaces will be elaborated in Section 4, below.

One important claim in the literature has been that any genetical system

can be adequately represented in at least two ways: its original, usually genotypic form, and a revised allelic one (Dugatkin and Reeve 1994; Sterelny 1996; Kerr and Godfrey-Smith 2002). In order to assess whether this is correct, we must address the overall issue of how to compare models in population genetics. We propose a new definition of representational adequacy that rests on dynamical and parametric sufficiency, from which equivalences can be derived.

3. Representational Adequacy of Models. The most common approach to comparing population genetic models has emphasized prediction of allele frequency changes: If two models both predict the same changes in allele frequencies, it is thought, then the models are equivalent. But this is an inadequate approach to understanding and confirming models, as will be demonstrated below. We advance the notion of *representational adequacy*, which we define as parametric and dynamical sufficiency. Why introduce representational adequacy? What is wrong with straightforwardly checking whether the model fits the allele frequency data?

There are a variety of ways that models can be tested against data, and fitting the outcome or prediction of the model—in these cases, the predicted allelic or genotypic frequencies—is only one of them (Lloyd 1987). Other significant components of the empirical evaluation of any mathematical model include: testing the values of its parameters against the system independently (e.g., measuring or estimating the mutation parameter value in the model); evaluating the appropriateness of the state space and parameters used; and testing the model against a range of values in the variety of systems to which it is supposed to apply (variety of fit). In addition, a model is taken to be better confirmed when it has more of its parameter values—that is, a variety of them—estimated or confirmed independently (Lloyd 1994, 145–159).

Our notion of representational adequacy combines the traditional standards of predictive accuracy and goodness of fit with the broader requirements of confirming that the state space, parameters, and laws being used in the models are appropriate and sufficient to the task (see Skipper 2004; Forber 2008). We take it as foundational to any notion of adequate representation that the standards of parametric sufficiency in model-building be weighed in judging overall model adequacy. Parametric sufficiency is dependent upon choice of space and parameter set, in any particular case.

The concept of dynamical sufficiency is precisely defined in terms of a set of objects and their frequencies, and another set of objects and their frequencies. Dynamical and parametric sufficiency together provide a much more adequate measure of a model's empirical worth than the vague notion of empirical adequacy, or the overly simplistic idea that if a model's

prediction of allelic or genotypic frequency is correct, then the model is empirically substantiated.

3.1. Parametric Sufficiency. Parameters are properties of objects, and may be properties of more than one object at a time. (Parameters are represented in the models as values that are not variables.) Two models may look as though they should be dynamically equivalent because of the similar appearance and *names* of their parameters, but real differences in their parameter measures may result in dynamics that are not equivalent. Two models should be considered 'parametrically equivalent' if the parameters that apply in one model have a natural representation in the parameters that apply in another. One case in which similar-looking parameters yield very different dynamics concerns allelic and genotypic fitnesses, which has led to much confusion in current controversies. (See Sections 5.1 and 5.3.)

3.2. Dynamical Sufficiency. The concept of dynamical sufficiency concerns what state space and variables are sufficient to describe the evolution of a system given the parameters being used in the specific system. What happens to the frequency of the variable over time? In a simple allelic model, this question becomes: Can we describe the changes in the frequency of allele A over time, with the information that we have, which includes the state space (variables) and the parameters (fitnesses, population size, etc.)?

Consider a set of objects $\Omega_1 = \{O_1, O_2, \ldots, O_m\}$ (e.g., genotypes $O_1 = AA$, $O_2 = Aa$, $O_3 = aa$) whose frequencies are given by $\kappa = \{k_1, k_2, \ldots, k_m\}$, and a set of variables $G = \{g_1, g_2, \ldots, g_n\}$ that represents the frequencies of another set of objects $\Omega_2 = \{O_1, O_2, \ldots, O_n\}$. For example, $\Omega_2 = \{A, a\}$. Then Ω_2 is *dynamically sufficient* for Ω_1 if the functions $\{f_i\}$ and $\{h_i\}$ are such that we can write the recursions

$$g_i^{t+1} = f_i\{g_1^t, g_2^t, \ldots, g_n^t\} \quad (i = 1, \ldots, n)$$

with

$$k_j^{t+1} = h_j\{g_1^t, \ldots, g_n^t\} \quad (j = 1, 2, \ldots, m).$$

In the one-locus diploid case with selection on AA, Aa, aa in terms of fitnesses w_{11}, w_{12}, w_{22} (stage 7 above) we have

$$\Omega_1 = \{AA, Aa, aa\}, \qquad \kappa = \{x, y, z\}$$

$$\Omega_2 = \{A, a\}, \qquad G = \{p_A, 1 - p_A\}.$$

Then after selection

$$g_1^{t+1} = p_A^{t+1} = \frac{(1 - p_A^t)(p_A^t w_{11} + (1 - p_A^t)w_{12})}{w_{11}(p_A^t)^2 + 2w_{12}p_A^t p_a^t + w_{22}(1 - p_A^t)^2},$$

$$\kappa_1^{t+1} = x^{t+1} = (p_A^{t+1})^2,$$

$$\kappa_2^{t+1} = y^{t+1} = 2p_A^{t+1}(1 - p_A^{t+1}),$$

$$\kappa_3^{t+1} = z^{t+1} = (1 - p_A^{t+1})^2.$$

In this case, the allele frequencies G in the allelic state space are dynamically sufficient to study the evolution of genotype frequencies κ. These equivalences depend, however, on the parameters, w_{11}, w_{12}, w_{22}, that could only be determined in the genotypic space.

Increasing attention has recently been paid to the phenomenon of epigenetics, which includes a variety of biological processes that act on genes and may be transmitted between generations, but not according to any of the rules of genetic inheritance. Formal models for the evolution of epigenetic objects or properties (Feldman and Cavalli-Sforza 1976, 1981; see also Jablonka and Lamb 2005) utilize an additional state space S_7, the phenogenotype, in which changes may occur during organismal development.

4. Population Genetics Models: Dynamical Issues. Returning to our discussion of the six state spaces involved in the calculation of a single generational change in genetics, we elaborate several of the steps involved in mapping from space to space, in order to illustrate the assumptions undertaken when laws are formulated and parameters are chosen, estimated, or omitted.

Philosophers have focused much attention on genotypic space, undoubtedly because genotypes are the 'smallest' entities that have dynamically adequate state spaces within which to calculate the next generation's allele frequencies (Section 5.1). But in order to do so adequately, genotype models must incorporate biological information—that is, the parameters—from other state spaces involved in the generational cycle. Start with the fact that we cannot calculate the next generation of genotypes without a segregation rule, that is, without knowing whether the genotypes undergo a normal process of meiotic distribution of gametes—information from the haploid model.

When we move from the diploid model to the haploid representation, in which only allelic frequencies p and q $(= 1 - p)$ are tracked, things get tricky. When philosophers and biologists speak of allelic space, they refer to the space of allelic frequencies. But in such a space, the marginal

fitnesses $w_{1.}$ and $w_{2.}$ are calculated from the diploid genotypic fitness parameters w_{11}, w_{12}, and w_{22}, without which they could not be evaluated or assigned (see Lloyd 2005; see also Section 5). Thus, the p's are in one (allelic) space, and the w's refer to another (genotypic) space. The state transition is actually calculated using parameters from genotypic space, and then translated or mapped back into allelic space. Allelic space plus allelic parameters alone are not able to represent the evolution of this genetic system. As the examples below will show, this is not merely the problem that marginal allelic fitnesses depend on genotypic parameters. In many important cases, generational transitions for the genotypes cannot be calculated from allele frequencies.

4.1. Example—Dynamical Insufficiency: Allelic Frequencies and Self-Fertilization. Again $\Omega_1 = \{AA, Aa, aa\}$, $\kappa = \{x, y, z\}$, $\Omega_2 = \{A, a\}$, $G = \{p_A, 1 - p_A\}$. With random mating and no selection, after the first generation,

$$x^t = (p_A^t)^2, \qquad y^t = 2p_A^t(1 - p_A^t), \qquad z^t = (1 - p_A^t)^2.$$

The allele frequencies are constant, $p^{t+1} = p^t$, $t \geq 0$, as are the genotype frequencies after the initial generation. With self-fertilization we have

$$\begin{aligned} x^{t+1} &= x^t + y^t/4, \\ y^{t+1} &= y^t/2, \\ z^{t+1} &= z^t + y^t/4. \end{aligned}$$

Again $p^{t+1} = x^{t+1} + y^{t+1}/2 = p^t$, and allele frequencies are again constant over time. But the heterozygote frequency y^t tends to zero at the rate 1/2 per generation. Thus, in the case of selfing, allele frequencies do not give the dynamical properties of the full system, which converges to a mixed population of AA and aa homozygotes with frequencies p, $1 - p$, respectively. Partial self-fertilization would have a similar qualitative difference between allelic and genotypic evolution.

The lesson from such cases is not, however, that genotypic space is itself somehow magically sufficient or transparently representative of the evolutionary system, for genotypic parameters, particularly the fitness parameter, are most often averages, estimated in a number of ways.

4.2. Fertility Selection. The issue of types mating with types arises similarly with fertility selection. Here the types of matings produce different numbers of offspring, which means that the fertility-level fitness parameters are properties of the mating pairs. Bodmer (1965) shows that it is not possible to represent the results of differential fertility among mating types in general in terms of allelic state space. Having given the equations

that relate the proportions of genotypes in one generation (right-hand side of the equations) to those same genotypes in the next generation (left-hand side of equations), he then remarks: "These equations are intractable as they stand, and allow no simple expression of the right-hand sides in terms of gene frequencies" (Bodmer 1965, 412). The population dynamic is then necessarily presented in genotypic space, not in terms of genic types. Most often, planktoid mating is simply assumed—that is, the random union of gametes—obviating the need for any additional parameters beyond fertilities of the mating pairs of genotypes.

4.3. Continuously Breeding Population. A much greater complication of prediction occurs when diploid organisms with overlapping generations and continuous breeding are considered. In addition to the age specific mortality and fecundity schedules of the different genotypes, it is essential to specify the growth rate of the population as a whole, particularly whether the population is expanding, stable in size or shrinking. For example, in a growing population, genotypes with reproduction at earlier ages increase in frequency since, relative to those that postpone reproduction, they are contributing their genotypes while the population is at a smaller size. The same genotypes will decrease in frequency in a shrinking population, since it is advantageous to wait until the population has decreased in total size before adding copies of the genotype (Charlesworth and Giesel 1972). Thus the complete age-specific mortality and fecundity schedules are insufficient to predict gene frequency change and there is no 'genotype fitness parameter' at all in this model; genotypic fitnesses change continuously over time. The only way to arrive at answers about the evolutionary state of the system is to grind through the mathematical dynamics of the model. Note that, in this model, the biological character of the genetic types does not change, however we choose to model them. The changes in frequencies occur because the types are part of a population that is either growing larger or smaller. There is no reason to think that this sort of context-dependence is unusual; rather, it seems to be the rule. In that case, parameters that describe the context must be integrated into genotypic fitnesses, with the result that separating the context and genotypic dynamics from a forteriori allelic evolution becomes impossible.

Moreover, dependence on population growth or decline is not the only kind of context-dependence that matters, in looking at genotypic fitnesses. Lewontin demonstrated that fitnesses in genetics, including genotypic fitnesses, depend on what other genotypes are present in the population at that time (1974). Thus, we must conclude that the fixed genotypic fitnesses that we are accustomed to seeing in population genetic models are at best approximations, as close to the truth as we can afford at a given level of mathematical precision.

5. Problematic Claims. One widely repeated set of claims has revolved around Dugatkin and Reeve's (1994) assessment of the mathematical equivalence of a wide variety of population genetic models (Sterelny 1996, 577; Sober and Wilson 1998, 57, 98–99; Sterelny and Griffiths 1999, 168–169, 172; Kerr and Godfrey-Smith 2002, 479, 508; Waters 2005, 312).

5.1. Dugatkin and Reeve's Formulation of Allelic Equivalence. Dugatkin and Reeve claim:

> A number of theoretical investigations (Alexander and Borgia 1978; Uyenoyama and Feldman 1980; Wilson 1980; Colwell 1981; Crow and Aoki 1982; Michod 1982; Wade 1985; Maynard Smith 1987; Queller 1992a, 1992b) have shown that the mathematics of the gene-, individual-, kin-, and new group-selection approaches are equivalent. . . . We will show . . . that this must be the case. (1994, 108)

They do this with a pair of inequalities (1994, 109). Dugatkin and Reeve claim that their inequality (2) "encompasses both broad-sense individual selection and any form of trait-group selection that one may care to envision." Moreover, they conclude, "*If broad-sense individual selection, genic selection, and trait-group selection all can be represented by a single condition based only on allele frequencies, then they cannot fundamentally differ from one another*" (1994, 109; their emphasis). But there are serious problems here. The first is that their inequalities (1) and (2) hide completely the causes of why the numbers of alleles change.

Their inequalities (1) and (2) state

$$\frac{\sum_i p_i' N_i'}{\sum_i N_i'} > \frac{\sum_i p_i N_i}{\sum_i N_i}, \tag{DR}$$

where p_i and p_i' are the frequencies of the allele in group i before and after selection, respectively, and N_i and N_i' are the corresponding group sizes. Of course, inequality (DR) is misleading and illustrates the importance of both dynamic and parametric sufficiency. Hidden in (DR) is a 'fitness' of p_i, namely, $N_i/\Sigma N_i$, but N_i changes to N_i' over a generation, which requires an additional dynamic relationship and assumptions on the rule of demographic change.

For a true haploid dynamic, we would usually write, for one locus with alleles A and a and p_i the frequency of A in population i,

$$\bar{w}^i p_i' = p_i w_1^i, \tag{1}$$

where $\bar{w}^i = p_i w_1^i + w_2^i(1 - p_i)$, and w_1^i and w_2^i represent the fitnesses of alleles A and a in population i, and the prime indicates the next generation.

For the diploid one-locus case, we again write

$$\bar{w}^i p_i' = p_i[w_{11}^i p_i + w_{12}^i(1 - p_i)], \tag{2}$$

where

$$\bar{w}^i = p_i^2 w_{11}^i + 2p_i(1 - p_i)w_{12}^i + (1 - p_i)^2 w_{22}^i$$

is the mean fitness in group i, and $w_{11}^i, w_{12}^i, w_{22}^i$ are the fitnesses (viabilities) of AA, Aa, aa in group i. As in equation (1) it may seem that we can write

$$\tilde{\bar{w}}^i p_i' = p_i \tilde{w}_1^i, \tag{3}$$

where $\tilde{\bar{w}}^i = p_i w_1^i + (1 - p_i)\, w_2^i$. But now we are dealing with haploid allelic parameters, $\tilde{w}_1^i, \tilde{w}_2^i$, for the fitnesses of A, a in population i. Of course, although these appear to be haploid parameters they are actually expressible as

$$\tilde{w}_1^i = p_i w_{11}^i + (1 - p_i)w_{12}^i, \tag{4a}$$

$$\tilde{w}_2^i = p_i w_{12}^i + (1 - p_i)w_{22}^i, \tag{4b}$$

which involve genotypic parameters (as well as allelic frequencies). In other words, the apparently purely-allelic parameters depend crucially on genotypic fitnesses—but these values are *completely hidden*. The assumption of Mendelian transmission is also made in deriving (3) (see also Godfrey-Smith and Lewontin 1993). For this two-allele diploid case, if meiotic drive were introduced, one more parameter, k, giving the probability that A is produced by Aa heterozygotes, would be necessary.

Thus, we note that in describing the algebra, many discussions of allelic models and their algebra obscure the origins of that algebra and all the information it contains and represents. A common move is to infer that the genic state space, and its basic entity, the allele, have a metaphysically fundamental and autonomous character (Sterelny 1996). This is both biologically and mathematically problematic. Biologically, the changes represented are dependent on all changes in the entire generational cycle reviewed above, represented in the variety of spaces and parameters. Mathematically, because genic space is dynamically insufficient for representing many system changes, and because in diploids there are no allelic transition laws within allelic space with allelic parameters, there is nothing autonomous about it; thus, it cannot support the metaphysical inferences based on its supposed autonomy (see Section 4 above and Lloyd 2005). But there is another account available which might be thought to avoid some of the above problems.

5.2. Kerr and Godfrey-Smith's Formulation of Allelic Equivalence. Echoing Dugatkin and Reeve's claims of equivalence, which they endorse, Kerr and Godfrey-Smith (2002) similarly appear to be claiming parametric sufficiency for allelic fitnesses in the diploid one-locus population genetic model, but in a different fashion. They introduce interchangeable sets of parameters $\{\alpha_i, \beta_i\}$ and $\{\pi_i, \phi_i\}$ to describe the expected numbers of alleles in groups of size n. In the standard diploid population genetic model with Mendelian segregation, groups are of size 2 and the relative fitnesses of AA, Aa, and aa are in Kerr and Godfrey-Smith's notation $\pi_0 = 2\beta_0 = w_{11}$, $\pi_1 = \alpha_1 + \beta_1 = w_{12}$, $\pi_2 = 2\alpha_2 = w_{22}$. An additional parameter, ϕ_1, analogous to k in the case of meiotic drive, completes the four-parameter specification. Note that in the Mendelian case, three parameters are necessary to describe the fitnesses of the genotypes. The statement by Kerr and Godfrey-Smith that "the fitnesses of, for example, allele A in a given genotypic context is given by the α terms" and therefore "the dynamics of the population genetic model can be redescribed using the fitnesses of alleles" (2002, 499), gives the impression that only two parameters, properties of alleles A and a are sufficient, and is misleading.

In their footnote 27 they admit that their use of the term allelic fitness departs from the standard marginal fitness of alleles (given by the $\tilde{w}_i$ above). But their "context-dependent fitness" requires that a parameter be assigned to an allele for each "genotypic environment" in which it appears. Thus, in terms of parametric sufficiency, there are no allelic properties that are independent of genotypic properties. In other words, in their formulation, allelic fitnesses are not parametrically sufficient whereas genotypic fitnesses are.

Higher-level properties must be specified even in the case of meiotic drive where, in addition to the three genotypic fitnesses of AA, Aa, and aa, a fourth parameter is needed to specify the chance, k, that Aa produces gamete A. Marginal fitnesses of alleles A and a can be legitimately computed, but they are not parametrically sufficient because they are 'stated' or 'defined' in terms of w_{AA}, w_{Aa}, w_{aa}, and k. Empirically, of course, it is these four biological parameters that can be measured, whereas Kerr and Godfrey-Smith's α_i and β_i confound genotypic and gametic contributions to fitness. It is certainly not parsimonious to introduce allelic properties of genotypes that are made up of combinations of genotypic properties (of genotypes) and allelic properties (of gametes), which together make up a parametrically sufficient set of property values for use in the dynamics of the evolutionary process.

5.3. Group Selection/Altruism's Claimed Reduction to Allelic Formulation. One of Dugatkin and Reeve's most important and widely-repeated claims has been that "new-group selection" approaches are mathemati-

cally equivalent to kin- and individual-selection models (Sterelny 1996, 577; Sober and Wilson 1998, 57, 98–99; Sterelny and Griffiths 1999, 168–169, 172; Kerr and Godfrey-Smith 2002, 479, 508; Waters 2005, 312). But Cavalli-Sforza and Feldman showed the contrary in 1978, and the paper cited by Dugatkin and Reeve in their own support by Uyenoyama and Feldman (1980) reflected on this very result.

Consider the following model in which γ is the fitness loss by 'doing altruism' and β measures the gain in fitness. In the case of parent-to-offspring altruism where AA is an altruist, Aa is an altruist with probability h, and aa is selfish, and the fitness loss and gain are additive, we have, writing x, y, z for AA, Aa, aa frequencies and p, q for allele frequencies of A, a, with primes indicating the next generation,

$$\bar{\phi}x' = p^2(1 - \gamma) + p\beta(x + hy/2), \tag{5a}$$

$$\bar{\phi}y' = 2pq(1 - h\gamma) + \beta(xq + hy/2), \tag{5b}$$

$$\bar{\phi}z' = (q^2 + \beta hyq/2), \tag{5c}$$

where $\bar{\phi}$ is the sum of the right sides. This is a case where if we add $x' + y'/2$ to get p', the frequency of A in the next generation, it cannot be expressed solely in terms of p (Cavalli-Sforza and Feldman 1978). The allele frequencies are not dynamically sufficient; genotype frequencies must be used.

Later, Uyenoyama and Feldman (1980) showed what the consequences of this dynamical insufficiency of allelic frequencies could be in generating different outcomes in the genotypic analysis from those obtained under the 'allelic sufficiency' assumption that $x = p^2$, $y = 2pq$, $z = q^2$ in 5a, b, c). Thus, in true kin selection modeling there is neither dynamical sufficiency of alleles nor parametric sufficiency, because we need both genotypic fitnesses and donor and recipient fitness (or group-context) information. In other words, these models involve an interaction between genotypes, and not just a simple downward effect on genic frequency or even on genotypic frequency. Thus, these early models show that evolutionary dynamics of allele frequencies that are derived by approximating genotype frequencies are not equivalent to the dynamics of the genotypes when they are correctly counted.

Thus, empirical differences between systems in nature are involved in the models, and underlie the reasons that higher-level fitness and other parameters are introduced into the models in the first place, for the sake of dynamical sufficiency unattainable through other means (see Sections 5.1–5.3). To emphasize this important result: the claim that no higher-

level parameters are needed for the empirical sufficiency of multi-level or "new group selection" models is untrue. The biology involved in these models demands the measurement of higher-than-individual-genotype values; specifically, the properties of genotypes must be measured within their genotypic context.

Thus, we conclude that there are fundamental differences between many of the models that have explicitly been claimed to be 'mathematically equivalent' or 'empirically equivalent'. These models are neither dynamically nor empirically equivalent, and thus are not equivalent in any meaningful way.

6. Conclusion. We emphasize the necessity of considering biological facts from the subcellular and allelic level all the way up to the population level, when evaluating the appropriateness and adequacy of population genetic models. We focus attention on the biological objects involved in population genetic models, for which the parameters are specified or derived. Whether a model includes or omits particular parameters can have decisive consequences for its *representational adequacy*, a notion that we define in terms of dynamical and parametric sufficiency. The demands of representational adequacy for population genetics often go beyond calculational adequacy of allele frequencies; when the biology is different in different evolutionary systems, we may need to use different population genetic models.

Knowing and incorporating the actual biology of the objects of the model is always relevant to a successful model. Generally, representational adequacy of models in many cases rests on making assumptions that involve dynamical and parametric sufficiency, the state space, and the detailed structure of the laws that—while they may be fulfilled in the normal course of events—cannot be taken for granted. Any given evolutionary system may be represented in any of the six spaces used in calculating its generational change. Without the information from the seven-stage cycle of spaces often needed for each generational genetic calculation, correct parameter values and the correct mathematical form of the law for a given space are simply not available, and the transition laws will not work.

In sum, through focusing on biological objects and their properties, we conclude that there are fundamental differences between many of the models that have been claimed to be mathematically equivalent. The models are neither dynamically nor empirically equivalent, and thus are not meaningfully equivalent in population genetics. We have not ruled out all instances of equivalence among models in all classes, only some of the more prominent claims. A pragmatic pluralism regarding which model to develop and pursue, given that we don't know ahead of time which model

will fit a given case, has much to recommend it, as Dugatkin and Reeve argue at length. It might also be fruitful to explore the extent to which some limited classes of models may be intertranslatable, as for example, equivalences between kin and group selection models (see Uyenoyama and Feldman 1980).

REFERENCES

Alexander, R., and G. Borgia (1978), "Group Selection, Altruism, and the Levels of Organization of Life", *Annual Review of Ecology and Systematics* 9: 449–475.

Bodmer, W. F. (1965), "Differential Fertility in Population Genetics Models", *Genetics* 51: 411–424.

Cavalli-Sforza, L. L., and M. W. Feldman (1978), "Darwinian Selection and 'Altruism'", *Theoretical Population Biology* 14: 268–280.

——— (1981), *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Charlesworth, B., and J. T. Giesel (1972), "Selection in Populations with Overlapping Generations; II. Relations between Gene Frequency and Demographic Variables", *American Naturalist* 106: 388–401.

Colwell, R. K. (1981), "Group Selection Is Implicated in the Evolution of Female-Biased Sex Ratios", *Nature* 290: 401–404.

Crow, J., and K. Aoki (1982), "Group Selection for a Polygenic Behavioral Trait: A Differential Proliferation Model", *Proceedings of the National Academy of Sciences of the United States of America* 79: 2628–2631.

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

Feldman, M. W., and L. L. Cavalli-Sforza (1976), "Cultural and Biological Evolutionary Processes, Selection for a Trait under Complex Transmission", *Theoretical Population Biology* 9: 239–259.

Forber, P. (2008), "On Biological Possibility and Confirmation". Unpublished manuscript.

Godfrey-Smith, P., and R. Lewontin (1993), "The Dimensions of Selection", *Philosophy of Science* 60: 375–395.

Jablonka, E., and M. J. Lamb (2005), *Evolution in Four Dimensions*. Cambridge, MA: MIT Press.

Kerr, B., and P. Godfrey-Smith (2002), "Individualist and Multi-level Perspectives on Selection in Structured Populations", *Biology and Philosophy* 17: 477–517.

Lewontin, R. C. (1974), *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Lloyd, E. A. (1987), "The Confirmation of Evolutionary and Ecological Models", *Biology and Philosophy* 2: 277–293.

——— (1994), *The Structure and Confirmation of Evolutionary Theory*. Princeton, NJ: Princeton University Press.

——— (2005), "Why the Gene Will Not Return", *Philosophy of Science* 72: 287–310.

Maynard Smith, J. (1987), "How to Model Evolution", in J. Dupré (ed.), *The Latest on the Best: Essays on Evolution and Optimality*. Cambridge, MA: MIT Press, 147–150.

Michod, R. (1982), "The Theory of Kin Selection", *Annual Review of Ecology and Systematics* 13: 23–55.

Queller, D. C. (1992a), "A General Model for Kin Selection", *Evolution* 46: 376–380.

——— (1992b), "Quantitative Genetics, Inclusive Fitness, and Group Selection", *American Naturalist* 139: 540–558.

Skipper, R. A., Jr. (2004), "Calibration of Laboratory Models in Population Genetics", *Perspectives on Science* 12: 369–393.

Sober, E., and D. S. Wilson (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Sterelny, K. (1996), "The Return of the Group", *Philosophy of Science* 63: 562–584.
Sterelny, K., and P. E. Griffiths (1999), *Sex and Death: An Introduction to Philosophy of Biology*. Chicago: University of Chicago Press.
Uyenoyama, M. K., and M. W. Feldman (1980), "Evolution of Altruism under Group Selection in Large and Small Populations in Fluctuating Environments", *Theoretical Population Biology* 17: 380–414.
Wade, M. J. (1985), "Soft Selection, Hard Selection, Kin Selection, and Group Selection", *American Naturalist* 125: 61–73.
Waters, C. K. (2005), "Why Genic and Multilevel Selection Theories Are Here to Stay", *Philosophy of Science* 72: 311–333.
Wilson, D. S. (1980), *The Natural Selection of Populations and Communities*. Menlo Park, CA: Benjamin/Cummings.