The Empirical Non-Equivalence of Genic and Genotypic Models of Selection:

A (Decisive) Refutation of Genic Selectionism and Pluralistic Genic

Selectionism

(6,120 words)

Robert N. Brandon
and
H. Frederik Nijhout

Duke University

June 15, 2005

Abstract. Genic selectionists (Williams 1966 and Dawkins 1976) defend the view that genes are the (unique) units of selection and that all evolutionary events can be adequately represented at the genic level. Pluralistic genic selectionists (Sterelny and Kitcher 1988, Waters 1991, Dawkins 1982) defend the weaker view that in many cases there are multiple equally adequate accounts of evolutionary events, but that always among the set of equally adequate representations will be one at the genic level. We describe a range of cases all involving stable equilibria actively maintained by selection. In these cases genotypic models correctly show that selection is active at the equilibrium point. In contrast the genic models have selection disappearing at equilibrium. For deterministic models this difference makes no difference. However, once drift is added in, the two sets of models diverge in their predicted evolutionary trajectories. Thus, contrary to received wisdom on this matter, the two sets of models are not empirically equivalent. Moreover, the genic models get the facts wrong.

Genic selectionists (Williams 1966 and Dawkins 1976) defend the view that genes are the (unique) units of selection and that all evolutionary events can be adequately represented at the genic level. Pluralistic genic selectionists (Sterelny and Kitcher 1988, Waters 1991, Dawkins 1982) defend the weaker view that in many cases there are multiple equally adequate accounts of evolutionary events but that always among the set of equally adequate representations will be one at the genic level. There have been many arguments against these views, for example (Wimsatt 1980, Brandon 1982, Sober and Lewontin 1982, Lloyd 1988), and one might be forgiven for thinking that genic selectionism had been thoroughly refuted. But, at least in the sociological sense of that term, it has not been. A, perhaps the, reason for this is that the refutations have primarily relied on philosophically contentious views on scientific explanation and causation—views their opponents have not been willing to accept. What both sides in this debate have accepted is that the genic and higher-level accounts are

empirically equivalent.[1] This paper will show that is not the case, that the two accounts give dramatically different, incompatible, predictions in a broad class of cases. The predictions are factually different and the genic models consistently get it wrong. Given that virtually all philosophers and scientists accept the position that scientific theories should agree with known facts, we will refute genic selectionism without resort to anything that is philosophically controversial.[2]

1. The Cases. Let us start with the case that has been most discussed in this literature, a case of heterozygote superiority. Let us suppose there is a single genetic locus with two alleles, *A* and *a*. Thus there are three genotypes, *AA, Aa*, and *aa*. By definition the heterozygote *Aa* is superior in fitness to the two homozygotes. In general the fitness of the two

---

[1] But see Brandon and Burian (1982, introduction to part II) and Godfrey-Smith and Lewontin (1993).

[2] We can, and will, precisely characterize the positions to which our criticisms apply. Briefly, they apply to any model that implies that selection disappears at certain sorts of equilibria, e.g. an equilibrium produced by heterozygote superiority. It is harder to precisely say who holds such positions. We are confident that the canonical figures of genic selectionism, Dawkins (1982) and Williams (1966), clearly adopt the position we are addressing. Sterelny and Kitcher (1988) can fairly be accused of holding this position, though the lack of clarity with which they characterize their views makes any definitive statement impossible. Waters (1991 and forthcoming) clearly does not hold the position that we show is empirically false; however in § 3 we will argue that mainstream genic selectionists would not find his particular account appealing. Finally there is the sort of pluralism defended by Kerr and Godfrey-Smith (2002). Our argument is not relevant to that. However, we cannot help but see the sort of pluralism described there as one between two different sorts of genotypic selection. If a model contains genotypic fitnesses and genotypic frequencies we, and we would suspect most practicing biologists, would call it a genotypic model.

homozygotes need not be equal, but for simplicity we will assume they are since nothing hinges on that assumption. The standard genotypic model normalizes the fitness of *Aa* at 1 and assigns the fitness of (1 – *s*) to the two homozygotes (where 1 ≥ *s* > 0). Although the value of s = 1 is a mathematical, and biological, possibility; for our purposes we cannot focus on that value since it is what Brandon has termed a value of *maximal fitness difference* (Brandon forthcoming). Fitness values are at the point of maximal fitness difference when some =1 and some =0 and there are no intermediary fitness values. Drift is impossible at a maximal fitness differential (MFD) point. Since we are going to be interested in the interplay of drift and selection we will need to give *s* some intermediary value. For now let us assume *s* = 0.5. This model predicts a stable equilibrium that will be reached in a number of generations (depending on the initial starting point, and population size). At this equilibrium the frequencies of the two alleles are both 50%.

The genic selectionist looks at the same situation and describes it differently. The two alleles, *A* and *a*, are in competition with each other. Their fitness depends on their allelic environment. *A* has a low fitness when it finds itself paired with another copy of itself, but has a high fitness when paired with *a*. Conversely for *a*. Selection in this case is frequency-dependent. For example if *a* starts at low frequency, then it will almost exclusively find itself in the favorable allelic environment of being paired with an *A*. Thus *a* would be assigned a high fitness value and *A* a low

value. This difference in fitness between *a* and *A* persists for how ever many generations it takes to reach equilibrium, but the fitness difference decreases as equilibrium is approached. After the equilibrium is reached both have equal fitness.

The genotypic model says that selection among the three genotypes occurs each generation, but that because of Mendelism the *Aa* × *Aa* matings produce offspring of the genotypes *AA*, *Aa*, and *aa* in the ratio of 1:2:1. Thus the stable equilibrium is the result of selection and Mendelism (or more generally the genetic and mating systems).   But the genic selectionist must say that once the equilibrium is reached selection no longer occurs because the fitnesses of the two alleles are the same. As we will see shortly, *s*, the selection differential, is a crucial parameter in models of selection and drift. The genotypic model has *s* = 0.5 in our example. This is true at equilibrium and at every other point in gene frequency space, i.e. fitness is frequency-independent. However, for the genic selectionist selection is frequency-dependent. At equilibrium all alleles have the same fitness, so *s* = 0.

But when there are two alleles at a locus and they are selectively neutral evolutionary genetics gives a very clear prediction. The frequencies of the two alleles will **drift**. If that selectively neutrality were frequency-independent, then one of the two alleles would drift to fixation.[3]

---

[3] Absolute fixation will not occur if there is mutation. Then we would say that approximately 50% of the population will move towards 100% *a* with some mutational variance remaining. Similarly for *A*.

It is important to note that population size is irrelevant to the qualitative prediction just described.[4] However, for the genic selectionist the allelic fitnesses are not frequency-independent, and so as the population drifts away from the equilibrium point genic selection will increase in magnitude and push the population back towards equilibrium. Thus both the genic and genotypic models predict the same equilibrium point, but, as we will see, they predict different trajectories towards that point. The genotypic model predicts a stable equilibrium. This prediction is based on selection *actively* maintaining the equilibrium. The genic model is forced to say that no selection occurs at equilibrium. But if no selection is occurring then the population will drift (see Brandon submitted). Once drift moves the population sufficiently, selection will tend to move it back towards equilibrium.

It is important to note that these two models make factually different predictions that are detectably different even over a few generations. The general method to be used would be one of comparing the likelihoods of the two hypotheses (selection occurring vs. no selection) given the observed data. No data set would be absolutely incompatible with either hypothesis, but many data sets would allow us to confidently pick one hypothesis over the other.[5]

---

[4] Drift dominates selection when $4Ns \ll 1$ (where N is the effective population size and s the selection differential, see Roughgarden 1979, 74-79). For neutral characters, or neutral alleles, $s = 0$.

[5] The well-known example of the sickle cell allele in human populations inhabiting malaria-infested regions is a good example of an equilibrium

At this point it will be helpful to put these qualitative points into a more quantitative framework. Selection dominates drift when the quantity $4Ns \gg 1$ (where N is the effective population size). Drift dominates selection when $4Ns \ll 1$. When $4Ns \approx 1$ then we can expect drift and selection to both have effects. Whether or not drift occurs in a population is an empirical matter of fact. Methods for differentiating selection and drift are becoming increasingly sophisticated. The point is that what might seem to be an unanswerable quibble between the genic and genotypic selectionists—namely, whether or not selection is occurring at equilibrium—is in fact a substantive issue that makes for big differences in evolutionary predictions. The genic selectionist's mistake of thinking $s = 0$ at equilibrium results in mistaken predictions about the future evolutionary trajectory of the population. In particular, the mistaken value given to s results in mistaken values for $4Ns$, and thus results in mistaken predictions about drift.

We will present a single numerical example to illustrate our point and then present the theory that generalizes it. Suppose the effective population size, N = 150. Let the genotypic s = 0.02. Now suppose the population is perturbed from equilibrium so that the frequency of A, p, is

---

actively maintained by selection. Cavalli-Sforza has studied variation in blood groups among small villages above the town of Parma in Italy and compared that pattern of variation to that among the larger towns around Parma. He found a much larger between-group variation in the smaller mountain villages than among the larger towns in the plains. This difference appears to be the result of drift. Another interesting human example of drift is the bit of mitochondrial DNA that all humans share. This surely is the result of drift.

0.51 while the frequency of *a, q,* is 0.49. This change in frequency does not affect the genotypic *s,* and $4Ns = 12$, which is in the range where selection should dominate. Thus we expect the population to move quickly back towards the 50:50 equilibrium. On the other hand, the genic selectionist assigns fitnesses in a frequency-dependent way. Since *a* is now more rare than *A* it is (slightly) fitter. Normalized, its fitness is 1 and the fitness of *A* is $1 - s^*$, where $s^* = 0.0004$. (To avoid confusion let *s* denote the genotypic selection coefficient and s^* the genic selection coefficient, and recognize that s^* is a function of s and the allele frequency [see below].) Thus the genic selectionist says that $4Ns^* = 0.24$, which is in the region where we expect drift and selection to both have effects, but drift should dominate.

The dramatic difference between the two models does not depend on small population size; rather it directly depends on the dramatic difference in the quantities $4Ns$ and $4Ns^*$ calculated from the two different models (because s^* is a function of both s and the gene frequency; see below). For the case of heterozygote superiority we can make the distinction between the genic and genotypic predictions general as follows. For the genic case, assume the fitness of the heterozygote to be 1 and that of both homozygotes to be (1-s). Then

Fitness of allele A = $\omega A = (2p^2 (1-s) + 2p(1-p)) / p$,

and

Fitness of allele a = $\omega a = (2(1-p)^2(1-s)+2p(1-p)) / (1-p)$.

The normalized fitness of A = $\omega A / \omega a$ and s* is therefore (1 – $\omega A / \omega a$),

therefore

$$s^* = 1 - ((2p^2 (1-s) + 2p(1-p)) / p) / ((2(1-p)^2(1-s)+2p(1-p)) / (1-p)).$$

This equation thus expresses s* in terms of s and p. The genic critical value

is 4Ns*=1, can be simplified from the above equation by some algebra to

be:

$$4Ns^* = 4Ns (2p - 1) / (1 + ps - s) = 1.$$

This is the genic equivalent of the genotypic critical value 4Ns=1 (which, as

can be seen, does not depend on p).

Genic and genotypic selection differ in their predictions of when drift and

when selection will occur. This is in the area of (N,s,p) parameter space

where the genotypic critical value is >1 (predicting selection) and the genic

critical value <1 (predicting drift). A graph of (N,s,p) space showing the

4Ns=1 and 4Ns*=1 surfaces is shown in Figure 1. The region between the

two surfaces is the region of parameter space in which genetic and

genotypic selection scenarios make different predictions about the likelihood of drift versus selection.[6]

In general, predictions will differ when:

$1/(4s) < N < (1 + ps - s) / (4s (2p - 1))$

or when

$1/(4N) < s < 1 / (1 - 4 N - p + 8Np)$

or when

$0.5 < p < (1 - s + 4Ns) / (8Ns - s).$

Thus, it would be a misunderstanding of the point being made here to think that the two models make factually different predictions that we can empirically differentiate only in the long run. At least for the right parameter values (of N, p, and s) the predicted difference is detectable in the short-run. In fact, the long-term predicted **outcomes** of the two-

---

[6] Since the significance of the quantity 4Ns is statistical it is an oversimplification to think if it as an absolute cut-off between state-space regions where drift dominates vs. regions where selection dominates.

models are (virtually) the same. They both predict a 50:50 ratio of the two

alleles. But the predicted trajectories are quite different since the genic

model has drift dominating the evolutionary process at (and even near)

the 50:50 point, and selection becomes strong only far from that point. In

contrast, the genotypic model has selection acting equally strongly at all

points in state space and thus allows only minor fluctuations around the

equilibrium point. To reiterate, the two models make factually different

predictions about evolutionary trajectories, and these differences are

empirically distinguishable, at least in the right parameter range.[7]

---

[7] One might wonder if the point being made is somehow an artifact of
comparing genotypic s and genic s*. It is not. We can make the same sort
of comparison between two different sorts of genotypic equilibria. At the
sort of equilibrium relevant to our discussion selection actively maintains
the equilibrium value. By that we mean that selection is acting at the
equilibrium point. There are other sorts of genotypic equilibria. For
instance suppose we have negative frequency-dependent selection acting
on the two homozygotes AA and aa and that the heterozygote is always
intermediate between the two homozygotes. (One can easily imagine
plausible causal stories that would instantiate this model—e.g., predators
preying on the most common phenotype.) In this situation we can assign
the Aa the (frequency-independent) fitness of 1. The fitness of AA is
described by a line that intersects the Aa line at $p = 0.5$ and has a slope of s.
(Where p, the frequency of A, forms the horizontal axis and fitness the
vertical). The fitness of aa is described by a line that intersects the Aa line
at $p = 0.5$ and has a slope of -s. In this model the maximal fitness
difference (which occurs at $p = 0$ and $p = 1$) is s. Thus it is calibrated to
compare with the genotypic model of overdominance we have been
discussing, where the maximal fitness difference is also s. But in this new
model the fitness of the three genotypes are all 1 at the equilibrium point
of $p = 0.5$. Population geneticists would normally describe this as a stable
equilibrium, just like the equilibrium produced by heterozygote
superiority. However the behavior of this model is not nearly as stable
around the equilibrium point because there is no selection at equilibrium.
A graph very similar to Figure 1 could be produced to show the difference
in behavior of these two models. To our knowledge this distinction has
not be recognized by population geneticists before. We describe this

But in more realistic models it is highly unlikely that the long-term predictions will be the same, and so the point we are making becomes even more damning of genic selectionism when we move to models that more closely mirror reality. For instance, Nijhout (2003 and forthcoming) and Rice (1998, 2004) have shown that multi-locus systems, where the phenotype is produced by realistic nonlinear developmental-genetic mechanisms, produce fitness landscapes that have long ridges instead of peaks. If one locus were allowed to drift then, because of nonlinear (=epistatic) interactions among loci, the selection regime at other loci would change. Loci that were previously highly correlated with the trait under selection (and thus themselves under selection) can become neutral, and *vice versa*, as the frequency of the focal locus changes. Thus the long-term predictions of the two models—one that has strong selection at the focal locus, the other that has weak or no selection at that same locus—would certainly diverge.

This dramatic divergence in predictions is not limited to cases of heterozygote superiority. Cases of heterozygote inferiority have the same result. A stable equilibrium is reached. At that equilibrium the genotypic model says that selection plus sexual reproduction maintain the

difference between the two sorts of equilibrium situations by saying that the first is one that is actively maintained by selection while the second is not actively maintained by selection. Thus our criticism of genic selectionism can be summed up as follows: genic selectionism is committed to describing equilibria that are actively maintained by selection as equilibria where selection is not active. But the behaviors of such equilibria differ and there is no room for pluralism here.

equilibrium (just as in the case of heterozygote superiority—see Godfrey-Smith and Lewontin 1993 for a detailed discussion of the population genetics of heterozygote inferiority). In contrast, the genic selectionist is forced to say that no selection is occurring at equilibrium, and thus has no resources for explaining or predicting the stability of that equilibrium point. As before neutral theory takes over when alleles are selectively neutral and drift is the predicted outcome.

Let us briefly discuss one final case, that is less simple than the above, but that is of much greater generality. The case is one from quantitative genetics. Quantitative genetics deals with traits, like height, that vary continuously, rather than discretely, and that are influenced by multiple genetic loci. The basic descriptive vocabulary of quantitative genetics includes the mean and variance of the distribution of some quantitative trait. Selection can take a number of forms, the primary ones being directional, stabilizing and disruptive. Directional selection occurs when one extreme of the distribution is favored, e.g., when taller organisms are fitter. Disruptive selection occurs when two (or more) points in the distribution are favored, e.g., highest fitness is associated with organisms 2cm tall and 7cm tall, all other heights being selected against. Finally, stabilizing selection occurs when one point in the distribution is favored (e.g., the fittest are 5cm tall). All three forms of selection acting on

quantitative traits can result in (more or less) stable equilibria.[8] (The reasons for the necessity of the more of less qualification are not directly relevant to the present discussion.) We will focus on stabilizing selection, which is quite common in nature (Endler 1986, Kingsolver et al. 2001). Imagine we have a trait, say height, with a normal distribution in a population. This distribution has mean $\overline{X}$ and variance σ^2. Now let us impose artificial selection on this population by only allowing organisms within 1/2 of a standard deviation ($1/2\ \sigma$) to reproduce. Let us assume that height in this population is influenced by a number of genetic loci and that it behaves in a way that we would expect from the experience of plant and animal breeders (i.e., no funny business). After a single generation we would expect $\overline{X}$ to remain unchanged, but σ^2 to decrease. Now we repeat (with smaller σ), and continue for a number of generations. What happens? At some point we either run out, or nearly run out, of selectable variation. Mutation should continue to introduce variation, and development is noisy, so we wouldn't really expect to get a population with zero variance in height. But we would certainly expect that maintaining this selection regime would maintain a stable equilibrium. Selection maintains the mean $\overline{X}$ and squeezes the distribution tightly around $\overline{X}$. Thus it is responsible not only for the mean value but also for

---

[8] In the case of directional selection, the equilibrium would be reached only when selectable variance (nearly) runs out. But this is the common experience of plant and animal breeders.

the (small) variance. It counteracts mutation and recombination among loci, which tend to increase σ^2.

Genic selectionist rarely, if ever, discuss cases of quantitative genetics so it is a bit difficult to know what they would say about this case. But again their impoverished conceptual repertoire would seem to force them to say genic selection is not occurring at this equilibrium. But if not, then drift is the expected outcome, contrary to what we observe.[9]

2. Analysis. The first two cases, heterozygote superiority and heterozygote inferiority, are theoretically well understood. The third case from quantitative genetics is much more difficult theoretically.[10] Nonetheless all three cases are ones where selection actively maintains an equilibrium. That is, were we to remove selection, the populations would drift in their state spaces. In the first two cases that state space could be described either in terms of allele frequencies or in terms of genotypic frequencies. The third case is quite different where the state space is described in terms of the mean and variance of the trait value distribution. Still in all three cases we need to invoke selection to explain stability, and have a basis for predicting future stability only if we can recognize that selection is acting in each of these cases.

---

[9] In general when two scientific theories give empirically different predictions the next step would be to check nature to see which, if either, was making accurate predictions. Notice that we have skipped that step. We know that the genotypic model is correct. What does that say about the status of genic selectionism as a serious scientific research program?
[10] See Barton and Turelli (1989), and Turelli and Barton (1990).

The genic selectionist lacks this basis. He or she, following Williams (1966, p. 59), defines the selection coefficient of an allele in a population as the weighted mean value of its fitness in its various genetic environments. According to Williams, the genetic environment of an allele includes both the other allele it is paired with at that locus as well as all the genes at all the other loci. In general, the definition of an allele's fitness is given as a function of its fitness in its various genetic environments weighted by the frequencies of these environments.

$$\overline{W} = \Sigma(P_i W^i)$$

(Where W^i is the fitness of the target allele in the ith genetic environment, and P_i is the relative frequency of that environment.) In standard population genetic models, like our first two examples, at equilibrium different alleles at the same locus will have identical mean fitness. Although the situation is much more complex in our quantitative genetics example; it is a reasonable first approximation to say that different alleles at a locus will have identical, or nearly identical, mean fitnesses here as well.

The upshot of all this is that in the range of cases we have considered selection is required to maintain equilibrium. However this selection does not penetrate to the allelic level (no differential reproduction of alleles occurs), thus the genic selectionists fails to recognize the selection that is occurring and is therefore unable to accurately predict the future

evolutionary trajectory of such populations. For the genic selectionist, this is not a good thing.

We can make this point more general. If one defines fitness in terms of evolutionary consequences and one defines evolutionary consequences in terms of **change** (change in gene frequencies in the case of population genetics, change in mean trait value in the case of quantitative genetics), then one is doomed to make the mistake described above. To make the logic of the above even more explicit, we are assuming that a difference in fitness is a necessary condition of natural selection. Starting with Fisher (1930) many have argued in favor of defining fitness in terms of evolutionary consequences (where evolutionary consequence is explicitly or implicitly defined in terms of change). Recently some have argued for a point even stronger than Fisher's, i.e., not only should fitness be defined in terms of evolutionary consequences, it must be so defined (see Walsh et al. 2002, and Matthen and Ariew 2002). The above results apply to any such approach, but here we will focus on genic selectionism.

The point made here is not based on some esoteric cutting edge biological discovery and so one might naturally wonder why it has not been recognized before. We will not try to definitively answer that question, but will briefly set out three speculative answers.

First, although the literature on this controversy is large, to our knowledge this paper is the first to bring drift into the discussion. So one

possible answer is that philosophers (and some biologists as well) largely fail to appreciate the impact of drift on the evolutionary process.

Second, many philosophers and biologists involved in this debate seem overly concerned with the formal properties of certain models while losing sight of the fact that these models are simply tools for scientific prediction and explanation. A carpenter who becomes overly concerned with the formal properties of hammers soon becomes unemployed.

Third, it seems plausible to us that many have tacitly accepted a Newtonian view of evolution (perhaps inspired be Sober 1984). In particular, many have tacitly accepted the idea that the Hardy-Weinberg Law is an analogue of Newton's principle of inertia. According to this view, the Hardy-Weinberg law gives the zero-force condition for populations, and when in this condition populations remain at rest. As one of us has argued elsewhere (Brandon submitted), this is seriously mistaken. The natural state of populations is random movement, anything other than that requires special explanation. But we cannot pursue that further here.

3. Genic Selectionist Response. If what we have presented above is correct then genic selectionism is no longer tenable. How might the genic selectionist respond? We can think of three responses, ranging from silly to interesting but wrong. We will start at the silly end.

First a genic selectionist might say that we have violated the rules of the game being played. That the game was one where for any selection story being told the genic selectionist could tell an equivalent story at the genic level.[11] Thus far none of the stories involved drift, so we have somehow broken the rules by making drift relevant. But that is silly. We haven't made drift relevant, nature has. And it is worth recalling the point we made earlier; in the cases we have discussed the genic selectionists are forced to say that the alternative alleles are selectively neutral. That being the case, population size is irrelevant to the occurrence of drift.

An important positive point can be made in response to the above. In all real populations selection and drift are intermixed. They are both part and parcel of a constitutive process of any evolutionary system, namely, the sampling process (see Brandon forthcoming, and Brandon submitted). Thus no theory of selection can claim to be empirically adequate if it excludes drift. It is our contention that the inclusion of drift shows how the genic selectionist goes wrong in describing situations of equilibria maintained by selection.

Second, a genic selectionist might try to explain the stability of equilibria by invoking counterfactual selection. That is, although selection is not occurring at equilibrium, it would occur were the population to move away from equilibrium. From the genic perspective that is certainly

---

[11] A number of people have thought that genic selectionists are merely playing an intellectual parlor game that is entirely parasitic on real science. Lloyd (forthcoming) makes this case in a compelling way.

true. For instance in the case of heterozygote superiority, were the population to move from the equilibrium point where the allelic fitnesses are identical, they would cease to be identical and selection would move the population back towards the equilibrium. However it is hard to see how non-actual processes can maintain an equilibrium. That is, if the population does in fact remain at equilibrium (as it would be expected to if the genotypic fitness differentials are large), then the genic selectionist is forced to say that selection is not actually occurring. (Even though it would occur were to population to mover from equilibrium.) Thus there is no basis to explain the past persistence, or to predict the future persistence, of the equilibrium. To put this point more succinctly: As pointed out in §1, the two models predict different evolutionary trajectories. The genic selectionist cannot avoid that inconvenient fact by appealing to counter-factual selection.

Finally there is a possible response that is not so easy to counter. This is not a response that either Dawkins or Williams would be comfortable with since they are both committed to defining allelic fitness in terms of the *mean* fitness of the allele across all of its genetic contexts. We will give a simple example to show why Dawkins and Williams are committed to mean allelic fitness. Then we will explore a suggestion by Ken Waters that, if correct, would allow the genic selectionist to say that selection is still occurring at the sort of equilibrium situations we have been considering.

Dawkins (1976, pp. 40ff) gives an example of choosing the best rower in crew. Our example is exactly analogous to his but has one advantage, which we will see later. It is from the less rarified world of three-legged racing. Contestants in such races come in pairs. They stand shoulder to shoulder, wrapping their interior arms around each other. Their interior legs are tied together. Hence the term 'three-legged'. Suppose we have ten people: Alma; Bert; Clarence; Dora; Eileen; Fred; Gerhart; Hans; Inez and Jacques. We are asked to determine the best three-legged racer of the bunch and send him or her on to the county fair. Notice that we have not been asked to determine the best three-legged racing team. That would be easy. There are 45 possible pairings of the 10 people. If that were our task, we would simple time each pair over the prescribed course and then compare their times. The fastest time would pick out the best team. But we have been asked to find the best individual three-legged racer. By now you may have noticed the analogy to diploid organisms. Alleles in diploid organisms come in pairs. And if one of them does well in that organism, the other does equally well. But the genic selectionist wants to assign fitness values to individual alleles. They do so in the same way we will pick our racer for the fair.[12]

---

[12]One might take this example as a *reductio* of Dawkins' position. After all, if we are interested in winning a three-legged race we really should be picking the best *team*. For example, a 6′ 5″ racer is unlikely to do well when paired with a 5′ 5″ racer, even though both are very good three-legged racers. But we are assuming the county fair has dictated the rules to us. That is, we are trying to give Dawkins' idea a fair hearing.

We tie Alma and Bert together and time them over our course. Then Clarence and Dora. And so on through all 45 pairs making sure every racer has sufficient rest between runs. The Alma-Bert time is assigned to both Alma and Bert. In the end Alma has had nine different partners and so she has nine times. We add them and divide by 9 to get her mean time. We do the same for everyone else. The person with the lowest mean time wins. This captures the ideas of both Dawkins and Williams. Although genes in a population occur in numerous genetic contexts selection always has an arithmetic mean effect of each allele in the population. And the mean effect is its fitness (see Dawkins 1976, pp.40 ff., Williams 1966, pp. 58 ff.).

This makes a certain amount of sense, but, as we have seen above, it has the consequence of making the *genic* models predictively inaccurate in cases of stable equilibria or stabilizing selection. Waters (1991) argues that the genic selectionist is not stuck with defining fitness in terms of the mean effect across all genetic contexts. So in the case of heterozygote superiority we can say that there are two allelic environments _A and _a (where '_A' stands for the allelic environment of having *A* as the allele on the homologous chromosome and '_a' is the environment where *a* is the other allele).[13] *A* does poorly in the _A environment and well in the _a environment. Similarly for *a*. This, Waters suggests, is just like selection in

_____

[13] Waters (1991) argues, unconvincingly we think, that this idea is really what Williams (1966) had in mind. Sterelny and Kitcher (1988) adopt Waters' view of allelic environments.

heterogeneous environments. Imagine a sessile haploid organism that reproduces asexually by producing small buds that are distributed by the wind. The physical environment over which they are distributed is heterogeneous; it is divided into wet areas (W) and dry (D). There are two haploid genotypes, G_1 and G_2. G_1 does well in W, but poorly in D. G_2 does well in D but poorly in W. What happens? Our expectation is that G_1 and G_2 will be maintained in the population in a stable equilibrium. And it is selection that maintains this equilibrium. Why not make the analogous case for alleles *A* and *a* in the case of heterozygote superiority?

First, Waters seems to understand the notion of environment in a particular way that makes the genic account look sensible, rather than as a real, measurable, manipulable, explanatory concept (for the outlines of that concept see Brandon 1990 chap. 2). Brandon has have defined a *homogeneous selective environment* as an area within which relative fitnesses are constant. In contrast, an area is *heterogeneous* if across it there is $G \times E$ interactions, i.e., relative fitnesses change. The wet/dry example above is indeed an example of environmental heterogeneity. We might use a simple spatial scale to measure heterogeneity, or a simple temporal scale or something more complex (see Brandon 1990, chap. 2). But a homogeneous environment is something that really exists in space-time. And within that region differential reproduction occurs in a (statistically) consistent way. This is what allows this notion to play the crucial theoretical role that it plays in contemporary evolutionary biology (see

Brandon and Antonovics 1996 for discussion of three examples). And this is what allows it to be experimentally measurable (Antonovics et al. 1988). Waters abstract construct lacks both these features.

Where is the _A environment? It cannot be the space-time worm of some particular organism, especially if that organism is heterozygote. Because then it would be the _a environment as well. It is incoherent that one and the same space-time region can be two different environments for the same selective process.[14] Are we looking at too high a level? Perhaps it is a part of some particular cell in an organism? But which cell? A somatic cell interacting with the external environment? A sex cell? This is clutching at straws.

But things get worse for this account. The primary problem is that differential reproduction *does not* occur in any of these contexts.[15] The fates of the two alleles at a locus in a diploid organism are tied together, just like Jacques and Inez in a three-legged race. Jacques may be exceedingly fast, and Inez stunningly slow; but when they are tied together they get the same time. The following metaphor may be helpful. A particular selective environment is an arena for competition. We can imagine one horse consistently beating another horse on a 1-mile dry track, but consistently

---

[14] There is nothing incoherent in the idea that the same region of spacetime could be two different selective environments for two quite different selective processes, for instance processes involving two different lineages (say host and parasite).

[15] Not that it is impossible for it to. There are genuine cases of within organism, or within cellular selection processes, e.g., meiotic drive (see Brandon 1990 Chap 3 for further examples and discussion). But the examples we are focused on are not such cases.

losing to the other on a 1.5 mile wet track. Those are two different arenas

for competition.   Similarly one genotype of a grass may do better than

another in one spot where the soil is contaminated with lead, but one

meter away, where the soil is uncontaminated the fitness order of the two

genotypes is reversed.[16] That is how we individuate selective

environments—in terms of consistency and change of relative fitness. And

fitness we ultimately measure in terms of reproductive success. Thus it

makes no sense to talk of a selective environment for an allele within a

diploid organism (unless, again, something like meiotic drive is going on).

And so we conclude that Waters' suggestion is not one that is really

available to the genic selectionist. The notion of selective environment it

relies on is ultimately incoherent. Thus genic selectionists, and genic

pluralists for that matter, are left with the coherent, but empirically

inadequate, ideas of Williams and Dawkins. And thus genic selectionism

and genic pluralism have been refuted.

4. Concluding thoughts. We have focused on genic selectionism in this

paper, but the argument here has a much broader scope.   We have seen

that any attempt to define fitness in terms of evolutionary change is

doomed to the same failure as genic selectionism. Any such attempt will

---

[16] Don't take this analogy too far. Nothing said here is meant to deny the important truth that organisms construct their environments and that environments can co-evolve with an evolving population. See Lewontin (1983), Brandon and Antonovics (1996), Brandon (2001) and Laland et al. (1996)

mistake strong selection for no selection in a range of cases. And once we decide to use the models to make genuine predictions, as opposed to treating them as objects for scholastic contemplation, this mistake will lead to empirically incorrect predictions. Thus, for instance, Fisherian fitness, is not just wrong-headed, it is empirically inadequate.

Some time ago Bill Wimsatt (1980) derided genic selectionism as mere book keeping. He was right. He had in mind a number of philosophical flaws in the genic selectionist approach. Genic selectionism is not only explanatory empty (Sober and Lewontin 1982, Brandon 1990); it is entirely dissociated from real science (Lloyd forthcoming). Here the focus has been on the more mundane. It is false.

References:

Antonovics, J., Ellstrand, N. C., and Brandon, R. N. (1988). Genetic variation and environmental variation: Expectations and Experiments. In *Plant Evolutionary Biology,* eds. L. D. Gottlieb and S. K. Jain, pp. 275-303. London: Chapman and Hall.

Barton, N. H. & Turelli, M. 1989 Evolutionary quantitative genetics: how little do we know? Ann. Rev. Genet. **23**, 337-370.

Brandon, R. N.: 1982, "The levels of selection," in P. Asquith and T. Nickles (eds.), *PSA 1982*, *Vol. 1*, Philosophy of Science Association, East Lansing, MI, pp. 315-323.

_____. (1990). *Adaptation and Environment.* Princeton: Princeton University Press.

_____. (2001). "Organism and Environment Revisited," in *Thinking about Evolution: Historical, Philosophical, and Political Perspectives* (ed. by R. Singh, D. Paul, C. Krimbas, and J. Beatty) Cambridge University Press (2001), 336-352.

_____. forthcoming, "The difference between selection and drift: A reply to Millstein." *Biology and Philosophy.*

_____.: submitted, "The Principle of Drift: The First Law of Biology".

Brandon, R. N. and Burian, R. (eds.)(1984). *Genes, Organisms, Populations: Controversies over the Units of Selection*, Bradford Books, MIT Press.

Brandon, R. N. and Antonovics, J. (1996). The coevolution of organism and environment [expanded version of Brandon and Antonovics 1995]. In R. Brandon, Concepts *and Methods in Evolutionary Biology.* Cambridge: Cambridge University Press.

Dawkins, R.: 1976, *The Selfish Gene,* Oxford University Press, Oxford.

_____., 1982, *The Extended Phenotype,* Oxford: Freeman.

Endler, J. A., 1986, *Natural Selection in the Wild,* Princeton: Princeton University Press.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press.

Godfrey-Smith, P., and Lewontin, R. C., 1993, "The dimensions of selection," *Philosophy of Science*, **60**: 373-395.

Kingsolver et al. 2001 *The American Naturalist* **157**(3): 245-261.

Laland, K. N., Odling-Smee, F. J. and Feldman, M. W. (1996). The evolutionary consequences of niche construction: A theoretical investigation using two-locus theory. *Journal of Evolutionary Biology* **9**: 293-316.

Lewontin, R. C. (1983). Gene, organism and environment. In *Evolution from Molecules to Men,* ed. D. S. Bendall, pp. 273-285. Cambridge: Cambridge University Press.

Lloyd, E. A.: 1988, *The Structure and Confirmation of Evolutionary Theory,* Greenwood Press, New York.

_____.: forthcoming. "Problems with Pluralism"

Matthen, M. and Ariew, A.: 2002, 'Two Ways of Thinking about Fitness and Natural Selection', *Journal of Philosophy* **99**, 55-83.

Nijhout, H.F.: 2003, "The Nature of Robustness in Development," *BioEssays*   **24**: 553-563.

Nijhout, H. F. (forthcoming), "Complex Traits: Genetics and Evolution".

Rice, S.H.: 1998, "The Evolution of Canalization and the Breaking of Von Baer's laws: Modeling the Evolution of Development with Epistasis," *Evolution* **52**: 647-645.

Rice, S.H.: 2004, *Evolutionary Theory*, Sinauer Associates, Sunderland MA.

Roughgarden, J.: 1979, *Theory of Population Genetics and Evolutionary Ecology: An Introduction,* Macmillan Publishing Co., New York.

Sober, E.: 1984, *The Nature of Selection,* MIT Press, Cambridge, MA.

Sober, E. and Lewontin, R. C.: 1982, "Artifact, Cause and Genic Selection," *Philosophy of Science* **49**: 157-180.

Sterelny, K. and Kitcher, P. 1988, "The Return of the Gene," *Journal of Philosophy* **85** (7): 339-361.

Turelli, M. & Barton, N. H. 1990, "Dynamics of polygenic characters under selection." Theor. Pop. Biol. **38**, 1-57.

Walsh, D., Lewens, T., and Ariew, A.: 2002, 'Trials of Life: Natural Selection and Random Drift', *Philosophy of Science* **69**, 452-473.

Waters, K. 1991, "Tempered Realism about the Force of Selection," *Philosophy of Science* **58**: 553-573.

Williams, G. C.: 1966, *Adaptation and Natural Selection,* Princeton University Press, Princeton.

Wimsatt, W. C.: 1980, 'Reductionistic research strategies and their biases in the units of selection controversy', in T. Nickles (ed.), *Scientific Discovery, Volume II, Historical and Scientific Case Studies,* Reidel, Dordrecht, pp. 213-259.
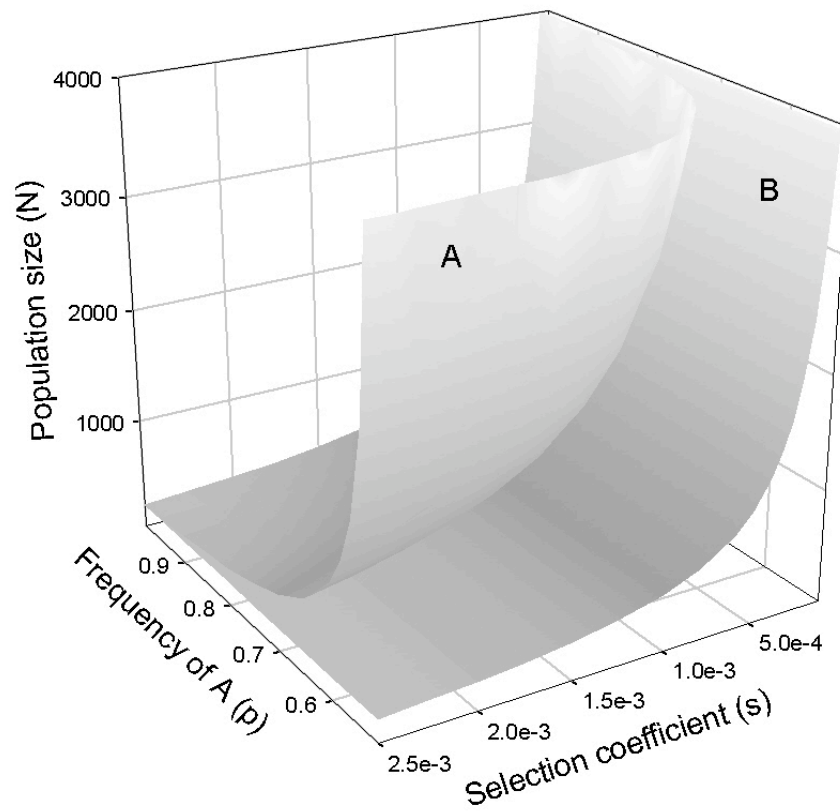
Figure 1. Surfaces that describe the critical condition 4Ns=1 for genic (A) and genotypic (B) selection, assuming heterozygote superiority. The region between the two surfaces is where the two models make different predictions about the relative importance of drift versus selection.