

Puzzles for ZFEL, McShea and Brandon's zero force evolutionary law

Martin Barrett · Hayley Clatterbuck ·
Michael Goldsby · Casey Helgeson · Brian McLoone ·
Trevor Pearce · Elliott Sober · Reuben Stern ·
Naftali Weinberger

Received: 30 November 2011 / Accepted: 27 April 2012 / Published online: 24 May 2012
© Springer Science+Business Media B.V. 2012

Abstract In their 2010 book, *Biology's First Law*, D. McShea and R. Brandon present a principle that they call “ZFEL,” the zero force evolutionary law. ZFEL says (roughly) that when there are no evolutionary forces acting on a population, the population's complexity (i.e., how diverse its member organisms are) will increase. Here we develop criticisms of ZFEL and describe a different law of evolution; it says that diversity and complexity do not change when there are no evolutionary causes.

Keywords Biological laws · Causation · Complexity · Drift · Force · Hardy–Weinberg · Stasis

In their provocative 2010 book, *Biology's First Law*, Daniel McShea and Robert Brandon develop an idea that they think is an analog of the law of inertia in Newtonian mechanics. Paralleling the idea in physics that an object remains at rest or in uniform rectilinear motion unless acted upon by a force, they (3–4)¹ propose a pair of ideas:

ZFEL (special formulation): In any evolutionary system in which there is variation and heredity, in the absence of natural selection, other forces, and constraints acting on diversity or complexity, diversity and complexity will increase on average.

ZFEL (general formulation): In any evolutionary system in which there is variation and heredity, there is a tendency for diversity and complexity to

¹ All such numbers refer to pages in McShea and Brandon (2010).

increase, one that is always present but may be opposed or augmented by natural selection, other forces, or constraints acting on diversity or complexity.

McShea and Brandon say that ZFEL (pronounced “zefel”) applies “in any and every dimension” (22), meaning that every trait in every population will increase in diversity and complexity (on average) if the condition specified in the antecedent of ZFEL (special formulation) is satisfied.²

We take it that the general formulation of ZFEL entails the special. If there is *always* a tendency for diversity and complexity to increase, then diversity and complexity *will* increase on average when there are no other causes. This is how the authors reason in their discussion of how ZFEL bears on the question of whether there has been an increase in complexity in the history of life—a tendency can be expected to manifest itself unless it is opposed by a stronger tendency (see especially p. 85). The fact that the general formulation entails the special means that any objection to the special formulation is also an objection to the general.

McShea and Brandon take pains to clarify what diversity and complexity should mean in ZFEL. These ideas pertain to variation—the more diversity there is among the parts of a whole, the more complex the whole is (10–11). So a population becomes more complex when variation among its member organisms increases. Although McShea and Brandon (135) decline to define variation in terms of the statistical concept of variance, they say that variance is often a “good measure” of how much variation there is. The authors argue that the right way to think about diversity and complexity is to strip these concepts of their functional significance (11). They believe that ZFEL should occasion a “gestalt switch” (128) that reorients our understanding of the evolutionary process and that reorganizes how practicing biologists should approach research problems (e.g., 130–133).

The present paper is not a thorough review of the rich range of ideas that McShea and Brandon develop. Rather, our more modest goal is to pose some puzzles for ZFEL. First we consider an objection to ZFEL to which the authors reply. We think their reply is unsatisfactory and that the objection is telling. Our second topic concerns what the authors mean by processes in different populations being “random with respect to each other.” Third, we consider what McShea and Brandon say about a rival to ZFEL, the Hardy–Weinberg Law in population genetics. Finally, we formulate a law of evolution that differs from ZFEL; it says that stasis is what happens when no causes are at work.

Puzzle 1: When variances are expected to decline in a drifting population

In a section called “An Apparent Anomaly Explained,” McShea and Brandon discuss the example of a population evolving when there are “absorbing barriers” (104–105). Consider a finite population in which there are two alleles (A and a) at a locus and the

² Without this “per trait” rider, ZFEL faces the problem of having to measure how diverse and complex a set of objects is overall, rather than just having to make sense of how diverse and complex they are with respect to this or that trait. This difficulty resembles a standard problem faced by pheneticism in its discussion of overall similarity.

gene frequencies are 50/50. If the two alleles are equal in fitness and there are no mutations, the allelic frequencies will do a random walk. As the number of generations of the process increases, the probability approaches unity that the population will evolve to 100 % A or to 100 % a. These two endpoints are “absorbing barriers” in the sense that once the population hits one of those end points, it is stuck there (there being no mutations to reintroduce the variation that has just been lost). If we use variance of gene frequencies as our measure of variation, then the population begins with the maximal degree of variation (given that there are just two alleles). There is nowhere for the variance to go but down. At the end of the process, the population has the lowest possible degree of variation (with respect to the locus in question).

For this to be a counterexample to ZFEL, the population must exhibit variation and heredity. It does, though the assumption of zero mutation means that no *new* variants can be introduced. For the example to be a counterexample, it also must be true that there are no forces or constraints that “act on variation and diversity.” By assumption, there is no selection, and the authors regard selection as a force. Drift is occurring in this population, but McShea and Brandon do not regard drift as a force; this is controversial, and will be discussed later. And let us assume that there is no migration, another vector quantity that the authors regard as a force. What about mutation? The authors, correctly in our view, regard mutation as a force (101) and cite Brandon (2006) as providing a “detailed argument” for thinking of it in this way. Mutation is like selection in that it has magnitude and direction and can change the genetic composition of a population.

If there is no mutation, selection, or migration in the population, isn’t the decline in variation that it experiences a counterexample to ZFEL? McShea and Brandon say *no*, on the grounds that the absorbing barrier is a “constraint on diversity” (30, 104). The problem is that this reply accomplishes too much. According to McShea and Brandon, when there are mutations, a force is present, and when there are no mutations, there is a constraint on variation. This means that a force or a constraint on variation must *always* be present. It follows that the antecedent of ZFEL (special formulation) is logically impossible. This is a hollow victory for ZFEL. No counterexample to ZFEL (special formulation) is possible because it is logically impossible for the antecedent of the following conditional to be satisfied:

In any evolutionary system in which there is variation and heredity, if there are no forces acting on diversity or complexity (so mutation is zero) and there are no constraints acting on diversity or complexity (so mutation is nonzero), then diversity and complexity will increase on average.

This is the special formulation of ZFEL, fleshed out by using what McShea and Brandon say about mutation and absorbing barriers.

There is nothing wrong with a zero force law’s having an antecedent that describes what will happen in situations that never, in fact, obtain (as in the law of inertia), but it is a step off the deep end for such “laws” to apply only to situations that are logically impossible.³ A conditional of the form “If X then Y” must be true

³ Could this problem be solved by distinguishing the level at which at which variation exists and the level at which the forces act? This won’t help with the present example, since McShea and Brandon correctly describe the variation at the level of allelic frequencies and the force of mutation as acting on allelic frequencies.

if X is logically impossible, no matter what proposition Y happens to be. This means that there are countless necessarily true conditionals that have the same antecedent as ZFEL, but have consequents that are incompatible with the one that appears in that principle. For example, consider the conditional that has same antecedent as the one displayed in the previous paragraph, but has as its consequent “diversity and complexity will *decline* on average.”

McShea and Brandon have two other replies to the example about absorbing barriers. They note that “in real populations mutations (not to mention migration) are always occurring” (104). This is true, but it does not remove the need to consider what ZFEL says about a situation in which mutations are absent. Consider the parallel point about the law of inertia; in our universe every object is acted on by a gravitational force, but the law says what will happen when all forces are absent. In assessing the law of inertia, it is relevant to consider what would happen were an object not acted on by a gravitational force. The authors also contend that the concepts of “allele” and “gene” that are used to describe the example about absorbing barriers are vestiges of a bygone age; they belong to the language of classical genetics, which possessed only very crude methods for defining the difference between one gene (or allele) and another. Molecular genetics has vastly improved our powers of discrimination, but with the result that allele is a “theoretical entity with no fixed molecular interpretation” (105). The authors take this to mean that once evolution is considered from the point of view of modern molecular biology, the problem of absorbing barriers can simply be swept aside. But it can’t. Take molecular units as small as you please (e.g., single nucleotides), and consider a model in which those units do a random walk with absorbing barriers. The problem persists.

Could this problem for ZFEL be solved by shifting from a single population that evolves in the way described to an ensemble of such populations? Suppose each of a thousand populations starts at 50/50 and that each does a random walk with no mutational input. If we wait long enough, we expect half the populations to evolve to 100 % A and the remainder to 100 % a. There was no variance among populations at the start and (in expectation) there is maximal variance among populations at the end. In this new, multi-populational, example, variation has increased, and so it is no threat to ZFEL. The problem is that ZFEL “predicts increase (in the absence of forces and constraints) at *every* single hierarchical level independently, regardless of how many there are” (47). ZFEL (special formulation) says that variability will increase “on average” (meaning in expectation) in each population and in the ensemble of populations as well, if forces are zero and constraints on variation are absent.

In view of the problem that arises for ZFEL when there are zero mutations, maybe ZFEL should be rewritten so that mutational input is assumed in the antecedent and the idea that mutation is a force is withdrawn. This newly-formulated principle is still vulnerable to a drift process that begins at 50/50. With very small mutational input, the population will, in expectation, evolve to *nearly* 100 % A or to *nearly* 100 % a (Durrett 2002, pp. 14–23). As before, the population begins with the variance at its maximum possible value and then the variance, in expectation, declines. The same pattern can arise when the population starts at a

frequency that differs from 50/50. If it starts at, say, 70 and 30 % and there is very small mutational input, drift will, in expectation, lead the variance to decline then too.

The problem we have described persists when there are more than two alleles at a locus. When there are n alleles and the starting frequency of each is $1/n$, the variability in the population cannot increase. And for other initial distributions of n alleles, the variability will probably decline as well, if the probability of mutating from one of those types to another is small.⁴

Perhaps McShea and Brandon will fault these examples for assuming a fixed stock of alleles. If mutation cannot enrich the stock of n allelic types, then it may be thought that this counts as a constraint on variation, and so ZFEL does not apply. But here again our argument does not rely on the assumption that the introduction of a new allelic type is impossible. Suppose there are n allelic types segregating in the population and their frequencies are about equal; if the probability is small, but not zero, that a new allelic type will be introduced by mutation, drift will probably have the result that the population's variability declines.

The lesson we draw from the example of a drifting population with absorbing boundaries is not that ZFEL is false; rather, we conclude that ZFEL, as stated, is not a law of nature, since it is logically impossible for its antecedent to be true. This result can be avoided by denying that mutation is a force and assuming that the populations to which ZFEL applies have variation, heredity, *and* mutational input. ZFEL now ceases to be vacuously true; it is false.

Puzzle 2: Random with-respect-to

The antecedent of the special formulation of ZFEL restricts that principle to cases in which there is no selection “acting on diversity.” This does not mean that the principle is restricted to cases in which there is no selection at all. For example, let's consider once again our thousand populations, but now suppose that all the populations have the same average adult height and that the selection processes that impact adult height in these different populations are “random with respect to each other” (17-19). The result is an increase in diversity among the populations. McShea and Brandon take this process to be an instance of ZFEL.

When the antecedent of ZFEL (special formulation) says that there is no selection “acting on diversity,” this does not mean that selection has no effect on diversity. Selection has precisely this effect in the example just described. What McShea and Brandon have in mind is that there is no selection *for* diversity, or *against* it.⁵ It is fine if there is selection for being tall in some populations and selection for being short in others. What is ruled out are examples like the following: imagine a metapopulation of groups, where variance in height within a group causes the group

⁴ When there are more than two allelic types in the population, it is natural to use the entropy, $-\sum p_i(\log p_i)$, of the probability distribution as a measure of variability. For n alleles, the entropy is maximal when all alleles have the same probability.

⁵ See Sober (1984, pp. 97–102) for discussion of the difference between selection-of and selection-for.

to do better or worse reproductively. Here there is group selection for higher (or lower) variance.

So when different populations experience selection processes that are “random with-respect-to each other,” this falls under the scope of ZFEL. What does this concept of randomness mean? Although this is not defined explicitly in the book, the examples that McShea and Brandon discuss provide some clues. The concept is meant to rule out cases in which two or more populations exert selection pressures on each other, as in arms races between predators and prey. And it also is intended to rule out cases in which several populations are impacted by a common cause that makes the populations converge (19). The randomness that the authors seem to have in mind occurs when the selection processes that impact the several populations are causally independent of each other. One way to spell this out is as follows: in the example we are considering, the thousand populations begin with the same average adult height. Then the selection pressures that the different populations experience are drawn independently from a flat probability distribution. For example, we might consider a range of optimal heights, say from 3 feet tall to 7 feet tall. For each population, a number is chosen uniformly at random from that range and the chosen number becomes the optimum towards which selection will push the population (in accordance with an Ornstein–Uhlenbeck process, on which see Lande 1976). If each of the thousand populations starts the process with an average height of, say, 5 feet, then the randomly chosen selection processes assigned to the populations will almost certainly cause the populations to diverge from each other.⁶ Even if the idea of drawing from a flat distribution isn’t necessary for what McShea and Brandon intend when they talk about selection processes that are “random with respect to each other,” surely it suffices, and that is enough for the point we want to make.

In the example just described of the thousand populations, a lot of work is being done by the initial conditions. If the populations begin with the same average adult height, of course there is nowhere for interpopulation diversity to go but up. Yet, there is no explicit reference to the initial state of the populations in either formulation of ZFEL. To see the relevance of this point, consider a third example. Now we have just two populations, one with an average adult height of 3’ 1”, the other with an average height of 6’ 11”. We choose optimal heights for the two populations from the same flat distribution described above. The result is that it is almost certain that selection will cause the two populations to converge, thus reducing interpopulation diversity. Notice that the selection pressures in the present example are drawn at random, just as they were in the previous example. What separates the two examples is their initial conditions.

Here is another example to consider. Let the average adult heights in our thousand populations differ from each other by as much or as little as you wish.

⁶ A similar scenario can be described for mutation. Again, imagine 1,000 identical populations but let each population’s mutation probabilities be drawn from a flat distribution. The forces of mutation at work in these different populations are then “random with respect to each other” and the populations can be expected to diverge. McShea and Brandon intend this process to be an instance of ZFEL, even though there is no selection and the process isn’t one of pure drift, either. This mutation process resembles McShea and Brandon’s example of the “evolution” of a picket fence (2–4); the pickets start in the same state and random insults then cause them to diverge.

Suppose the minimum average height exhibited by these populations is x and the maximum is y . Now let us draw, uniformly and at random, an optimum for each of these populations from a range of values that contains both x and y and is much larger; its minimum is much less than x and its maximum is much greater than y . The expected result is that the populations will diverge from each other. But notice that, once again, the expectation of divergence depends on the range of average heights that the populations exhibit at the start; the range of initial states must be smaller than the range of optima from which one is sampling.

Could this dependence on initial conditions be avoided if the optima were drawn uniformly and at random from a line that runs from zero to positive infinity (or, equivalently, on a log scale that runs from negative to positive infinity)? In this case, won't the populations almost certainly diverge from each other, regardless of how similar or different they are to each other at the start? Now we have gone too far. There is no such thing as sampling uniformly at random from this infinite range. On the other hand, there is nothing wrong with the idea of a *nonuniform* distribution over this infinite interval. However, if optima are chosen from this sort of distribution, there is no assurance that the populations will tend to diverge; that depends on what the nonflat distribution is. And if the optima are drawn at random from a finite range, we have already seen that there is no unconditional guarantee that the populations will tend to diverge.

Has our argument covertly assumed that there are constraints on how much variation the 1,000 populations exhibit at the start of the processes we have considered? Not at all. Choose any average heights you please for the populations' initial states. Whether they will then diverge from each other when selection occurs "at random" will depend on the distribution from which you select optima for each population.

We are not sure that the probabilistic reading of "random with-respect-to" fully captures what McShea and Brandon intend, though we think it would be strange if they were to deny that our formulation of drawing optima from a wide flat distribution falls under the category of selection processes that are "random with respect to each other." Sometimes McShea and Brandon seem to have in mind merely that different populations experience different selection pressures and that this will be enough for a selection-mediated ZFEL to cause diversity to increase, or that this will put in place a "tendency" for diversity to increase (18). However, mere difference in selection pressures is not enough. Returning to our two populations, one tall and the other short, suppose that selection acts on the 6' 11" population and pushes it towards an optimum of 6 feet, whereas selection on the 3' 1" population pushes it towards an optimum of 4. This set-up entails that the populations will tend to converge.

The two versions of ZFEL say nothing about the initial conditions that govern a population. Indeed, McShea and Brandon (13) say that ZFEL "is true for all starting conditions." But the examples described in this section show that the initial conditions matter when two or more populations experience selection pressures that are random with respect to each other.⁷

⁷ Our argument can also be applied to cases in which optima for populations are drawn from a normal distribution. This distribution can run from negative to positive infinity on a log scale. Whether optima drawn from a given normal distribution will cause the populations to converge or diverge depends on the variance in the normal distribution and on how much variance there is among the populations' starting trait values.

Puzzle 3: ZFEL v. the Hardy–Weinberg law

The main thesis of McShea and Brandon's book is that ZFEL is *the* zero-force evolutionary law. The definite article indicates that they think ZFEL is the real thing and that rival principles are unworthy pretenders. The one rival that McShea and Brandon discuss is the Hardy–Weinberg Law of population genetics. Just as they give two formulations of ZFEL, they do the same for Hardy–Weinberg (100–101):

(H-W₁) If a population exists with two alleles, A₁ and A₂, with frequencies p and q respectively, then in a single generation the population will settle into genic and genotypic equilibrium with gene frequencies p and q and genotypic frequencies A₁A₁ = p², A₁A₂ = 2pq, and A₂A₂ = q², provided that there is no selection, mutation, migration, nonrandom mating, or *drift* [their italics].

(H-W₂) If an infinite population exists with two alleles, A₁ and A₂, with frequencies p and q respectively, then in a single generation the population will settle into genic and genotypic equilibrium with gene frequencies p and q and genotypic frequencies A₁A₁ = p², A₁A₂ = 2pq, and A₂A₂ = q², provided that there is no selection, mutation, migration, or nonrandom mating.

McShea and Brandon think these principles have different defects. We think the two formulations are mathematically interderivable. For example, when selection and drift both occur in a population, the probability that the fitter of two traits will increase in frequency goes up as the product Ns increases (where N is the effective population size and s is the selection coefficient). This means that as N increases (and s is held fixed) the probability approaches unity that the fitter trait will increase in frequency. As N increases, the probability declines that the resulting change in trait frequency will deviate from its expected value. The probability of this happening is zero when the population is infinite. This is why “infinite population size” and “zero drift” are two ways of talking about the same thing.

McShea and Brandon say that (H-W₁) is “problematic as a zero-force law because it mixes genuine evolutionary forces ... with a nonforce, namely drift.” For this reason they think that (H-W₁) is “entirely misleading even if, strictly speaking, true (101).” Why isn't drift a force? “The basic argument,” the authors say, is that a force must be “a vector quantity and that each quantity can be measured with a common metric (e.g., genotypic frequencies) ... [B]ecause drift is not directional, it cannot be considered an evolutionary force.” The authors do not reply to an idea, presented by Sober (1984) and Stephens (2004), that drift has a direction—it tends to reduce heterozygosity. Perhaps they pass over this suggestion because Brandon (2006, p. 325) has already addressed it. Brandon considers a population with two alleles (A and a) whose starting frequencies are 50 % A and 50 % a. If there is drift but no selection, mutation, or migration, “we can predict that one of the two alleles will go to fixation.” Brandon says that “this is a prediction, but a prediction without a direction.” Filler (2011) replies that Brandon's assessment is correct for this specific example, but that the general conclusion about drift does not follow. For example, if the population had begun with a frequency of 70 % A, then its probability of evolving to 100 % A would be greater than its probability of evolving

to 100 % a. Notice that this reply to Brandon (2006) has to do with the expected loss of allelic diversity in a population. True, if the population is diploid, this will mean a decline in heterozygosity;⁸ however, this is incidental to the more general point, which also applies to haploid organisms.

Turning to (H-W₂), McShea and Brandon say that the problem here is that, “on the face of it,” (H-W₂) “applies only to infinite populations.” How, then, can this law apply to real world populations that are always finite? “One answer,” they say (101), “is that we simply stick drift back into the list of evolutionary forces, thus reverting to (H-W₁).” Notice that if (H-W₁) and (H-W₂) really are different principles, as McShea and Brandon say, this answer does not save (H-W₂), but merely changes the subject. McShea and Brandon then turn to a second answer, which is that (H-W₂) is a good approximation of real world systems and therefore can be used to make predictions. McShea and Brandon think this interpretation obliges one to take (H-W₂) to assert the following: “if no evolutionary forces act on a population, its gene and genotype frequencies will remain unchanged...” The authors then say: “But we know that [this is] false!” (102). The conditional is false, they believe, because ZFEL is true and ZFEL says that what will happen when there are zero forces is change, not stasis. This objection to (H-W₂) does not work; the principle does not use the *f*-word and it is not false (even if it does mix forces and nonforces).

McShea and Brandon have another reason to deny that the Hardy–Weinberg law is a zero-force law. They follow Beatty (1981) and deny that it is a law; it is “much more of an accidental generalization (100).”⁹ Their reason for denying that Hardy–Weinberg is a law is that it “depends on derived evolutionary conditions, diploidy and sexual reproduction, and so is not even true throughout the history of life on this planet, much less a part of universal biology.” This contrasts unfavorably with ZFEL, which McShea and Brandon think is part of “universal biology (100).” Here McShea and Brandon confuse the scope of the Hardy–Weinberg principle and the question of whether it is a law. True, Mendelian inheritance and diploidy appeared at a certain point in life’s history, and did so for contingent evolutionary reasons. But this does not mean that the Hardy–Weinberg principle isn’t a law. A conditional can be nomologically necessary even if its antecedent is not. Even if it is a contingent fact that there is gold in our universe, this leaves it open that “gold melts at 1,064.18 °C” may be a law.

There is a third formulation of the Hardy–Weinberg Law, one that McShea and Brandon do not consider:

(H-W₃) If a population exists with two alleles, A₁ and A₂, with frequencies *p* and *q* respectively, then in a single generation the population will (on average) settle into genic and genotypic equilibrium with gene frequencies *p* and *q* and

⁸ A defender of the gene’s eye view of evolution might insist that evolutionary forces must be defined solely in terms of their impact on gene frequencies, not in terms of their impact on genotype frequencies. McShea and Brandon do not take this position, which is no surprise, given that Brandon has been a strong proponent of the need to look beyond gene frequencies and genic fitnesses, and to include genotype frequencies and genotypic fitnesses, as well as other “higher-level” descriptors, in adequate models of the evolutionary process; see, for example, Brandon and Nijhout (2006) and Weinberger’s (2011) reply.

⁹ Beatty (1995) develops this point more fully and Sober (1997) replies.

genotypic frequencies $A_1A_1 = p^2$, $A_1A_2 = 2pq$, and $A_2A_2 = q^2$, provided that there is no selection, mutation, migration, or nonrandom mating.

This is just (H-W₁) with “on average” inserted and “drift” deleted. (H-W₃) resembles ZFEL (special formulation) in its logical form; in both cases, “on average” means *in expectation*. (H-W₃) says, in the circumstances described, that the expected outcome is stasis of allelic frequencies and stasis (after one generation) of genotypic frequencies.¹⁰

Still, the point remains that the Hardy–Weinberg law is part of Mendelian genetics and so it does not cover other possible mechanisms of heredity that might evolve in other galaxies, nor does it cover nongenetic modes of inheritance (e.g., epigenetic inheritance and cultural transmission) that are at work here at home. We agree with McShea and Brandon’s comment about this principle’s “evolutionary narrowness” (102).¹¹ ZFEL is intended to be more general, but it has the problems we described earlier. Perhaps, then, we should conclude that Hardy–Weinberg is the zero-force law of population genetics, but seek to find a replacement for ZFEL that can do a better job of furnishing a zero-force law in a universal biology.

Puzzle 4: Stasis redux

McShea and Brandon want ZFEL to apply to any population of objects that possess reproduction and heredity (21) and to do so without commitment to Mendelism’s being the mechanism that mediates inheritance. In this spirit, let us adopt the idea of (narrow sense) heritability, which simply means that the traits of parents and offspring are positively correlated; this is a good explication of what Darwin is getting at in the *Origin* when he says that offspring “tend” to resemble their parents. For simplicity, let’s consider the case in which there is uniparental inheritance in a population and each organism has one of two phenotypic traits (A and B). Each trait has a fitness (which may or may not change as the population evolves) and we will measure the intensity of selection by a selection coefficient, which is simply the variance in trait fitnesses. The less variance in fitness, the less selection there is. Just as selection has its intensity, so does drift, which we measure by the population size. The larger the population, the greater the probability is that the future state of the population, given its present state, will be close to its expected value. Just as the intensity of selection diminishes as variation in fitness declines, so the intensity of drift diminishes the larger the population is. Here we are talking about selection and drift as causes, not as outcomes (Sober 1984; Millstein 2002; Stephens 2004; Reisman and Forber 2005;

¹⁰ This third formulation of Hardy–Weinberg is perhaps a good place to answer McShea and Brandon’s question about how the law applies to real populations. The law is often applied in population genetics by using its contrapositive form: if a population deviates (significantly) from Hardy–Weinberg frequencies, biologists conclude that an evolutionary cause or force is present. We also want to emphasize that Hardy–Weinberg describes a sufficient condition for a population to have expected genotype frequencies of p^2 , $2pq$, and q^2 . The conditions described are not necessary; for example, a population can exhibit these frequencies when there is stabilizing selection.

¹¹ The definition of evolution as change in the genetic composition of populations is defective for the same reason (Sober 1993, p. 5).

Shapiro and Sober 2007). These ideas are summarized in the accompanying 2-by-2 table. All of what we have just said about the representation of traits A and B accords with what McShea and Brandon say; they are okay with the idea that selection and drift *cause* changes of trait frequencies in populations (105), though, as discussed above, they think that only one of these is a “force.”

	Population size	
	Finite	Infinite
Selection coefficient		
Nonzero	Drift and selection	Selection and no drift
Zero	Drift and no selection	No selection and no drift

If there is no selection and no drift (and no mutation or migration, either), then trait frequencies will remain the same; this is a “law of stasis.” The truth of this law does not depend on whether you regard drift as a force; notice that the *f*-word does not appear in our statement of this principle. True, no real population ever satisfies the antecedent of this conditional; real populations are always finite, mutations occur, some traits vary in fitness, and so on. But that does not affect the correctness of this law (anymore than the fact that all the objects in our universe experience a gravitational force threatens the status of the law of inertia). Whether one calls this a “zero cause evolutionary law” (a ZCEL) or a ZFEL is a secondary matter.

In our section on Puzzle 1, we held against ZFEL the fact that its antecedent is impossible. If there is no mutation in a population, this counts as a constraint on diversity; if there is mutation, this counts as a force. So no possible population can have no constraints on diversity and no forces acting on it. Does the same complaint apply to the principle we have just described? Let us suppose that it is impossible for a population to be infinitely large. The reason this is not a problem for the principle we have described is that talk of infinite populations can be understood in terms of the idea of a mathematical limit. Fair coins can’t be tossed infinitely many times, but it still is true, for any $\epsilon > 0$, that as the number of tosses is increased, the probability approaches unity that the frequency of heads will be within ϵ of 0.5. Talk of infinite sample size is just a way of describing this limit. This is the harmless sense in which there is no drift when population size is infinite. The fact that real coin tosses and real populations can’t, in fact, be infinite is irrelevant.

Concluding comments

Does our “law of stasis” disagree with McShea and Brandon’s ZFEL? We did not use the word “force” in stating this principle, and the two principles describe the same inventory of causes. So McShea and Brandon should agree that when the causes listed are zero, there will be no change. One issue that remains is whether

drift is a force. But there is more. ZFEL has the two flaws described in our sections on Puzzles 1 and 2. It remains to be seen whether these defects can be repaired.

We want to emphasize that we agree with McShea and Brandon about some of the examples they discuss. First, there is the case of drift: consider the average adult height exhibited by each of a set of populations. If these populations do a random walk on a line that runs from zero to positive infinity, the populations will in expectation become more separated from each other. This is true regardless of how similar in height they are at the start. And if a set of populations begins with each exhibiting the same average adult height, selection pressures drawn uniformly at random from a finite interval will in expectation make those populations diverge from each other. The same is true if those initially identical populations have mutation probabilities that are sampled uniformly at random. But ZFEL goes way beyond these examples; it isn't limited to cases in which two or more populations experience pure drift or all have the same initial state.

We also agree that zero-force (or zero-cause) laws can be theoretically important. Newton's law of inertia was conceptually central to his physics. A zero law for evolutionary theory might occupy the same central place. As the transition from Aristotle to Galileo and Newton makes clear, it is not trivial to say what will happen when no causes impinge. Aristotle thought that objects stop moving when no cause is present. Galileo and Newton showed that this is wrong; what will happen is constant velocity (i.e., zero acceleration).

When a science provides different models for what will happen when there are different combinations of causes, it is useful to organize those models by starting with a zero-force (or zero-cause) law. Next come the singleton laws, which describe what will happen when just one cause is present; then there are 2-force models, 3-force models, and so on. The models tend to become more realistic and more applicable as they become more complex. This way of structuring the science makes it seem less like a jumble of models and more like a coherent, systematic unity (Sober 1984).

Newton's law of inertia and the more general claim that $f = ma$ are able to have this centrality in physics partly because they are quantitative. The amount of acceleration an object of given mass exhibits tells you how much net force the object experiences. Your task, then, is to find the component forces that produced that net. Brandon and McShea (130) are right that it is an important task for the future that ZFEL be given a quantitative formulation. *How much* should diversity increase if zero forces impinge? If that baseline value were specified, an investigator could learn about the forces that are present by comparing the observed change in diversity to the baseline. But first ZFEL must be revised to take account of the problems we have described. If this cannot be done, then perhaps ZCEL should replace ZFEL. This is not to say that ZCEL has the practical benefits for working scientists that McShea and Brandon claim for ZFEL, but ZCEL at least has the virtue of avoiding the problems described in this paper to which ZFEL falls prey.

Acknowledgments We are grateful to Dan McShea and Robert Brandon for their helpful comments on an earlier draft, and to the anonymous referees of this journal for theirs.

References

- Beatty J (1981) What's wrong with the received view of evolutionary theory? In: Asquith PD, Giere RN (eds) *PSA 1980, vol 2*. Philosophy of Science Association, East Lansing, MI, pp 397–426
- Beatty J (1995) The evolutionary contingency thesis. In: Wolters G, Lennox JG (eds) *Concepts, theories, and rationality in the biological sciences*. University of Pittsburgh Press, Pittsburgh, pp 45–81. Reprinted in Sober E (ed), *Conceptual issues in evolutionary biology*, 3rd edn. MIT Press, Cambridge, 2006
- Brandon R (2006) The principle of drift: biology's first law. *J Philos* 103:319–335
- Brandon R, Nijhout H (2006) The empirical nonequivalence of genic and genotypic models of selection: a (decisive) refutation of genic selectionism and pluralistic gene selectionism. *Philos Sci* 73:277–297
- Durrett R (2002) *Probability models for DNA sequence evolution*. Springer, New York
- Filler J (2011) *Essays on explanation and evolution*. Ph.D. Dissertation, Philosophy Department, University of Wisconsin, Madison
- Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334
- McShea D, Brandon R (2010) *Biology's first law*. University of Chicago Press, Chicago
- Millstein R (2002) Are random drift and natural selection conceptually distinct? *Biol Philos* 171:33–53
- Reisman K, Forber P (2005) Manipulation and the causes of evolution. *Philos Sci* 72:1115–1125
- Shapiro L, Sober E (2007) Epiphenomenalism—the do's and the don'ts. In: Wolters G, Machamer P (eds) *Studies in causality: historical and contemporary*. University of Pittsburgh Press, Pittsburgh, pp 235–264
- Sober E (1984) *The nature of selection: evolutionary theory in philosophical focus*. MIT Press, Cambridge
- Sober E (1993) *Philosophy of biology*. Westview Press, Boulder
- Sober E (1997) Two outbreaks of lawlessness in recent philosophy of biology. *Philos Sci* 64:S458–S467. Reprinted in Sober E (ed), *Conceptual issues in evolutionary biology*, 3rd edn. MIT Press, Cambridge, 2006
- Stephens C (2004) Selection, drift, and the 'forces' of selection. *Philos Sci* 14:459–470
- Weinberger N (2011) Is there an empirical disagreement between genic and genotypic selection models? A response to Brandon and Nijhout (2006). *Philos Sci* 78:225–237