European Journal for Philosophy of Science (2021) 11: 112
https://doi.org/10.1007/s13194-021-00430-0

PAPER IN THE PHILOSOPHY OF THE LIFE SCIENCES

Fitness: static or dynamic?

Peter Takacs[1] · **Pierrick Bourrat**[2]

Received: 10 May 2021 / Accepted: 10 November 2021 /

Abstract

The most consistent definition of fitness makes it a static property of organisms. However, this is not how fitness is used in many evolutionary models. In those models, fitness is permitted to vary with an organism's circumstances. According to this second conception, fitness is dynamic. There is consequently tension between these two conceptions of fitness. One recently proposed solution suggests resorting to conditional properties. We argue, however, that this solution is unsatisfactory. Using a very simple model, we show that it can lead to incompatible fitness values and indecision about whether selection actually occurs.

Keywords Fitness · Natural selection · Evolutionary theory · Explanation · Causation

1 Introduction and background

Actual lifetime reproductive contribution or *realized fitness* may appear to be the most basic measure of organismal fitness (Rosenberg & Bouchard, 2015). Although this so-called "naïve view" is appealing in its simplicity, biologists and philosophers alike have successfully shown that actual survival and reproductive output cannot suffice as a direct measure of organismal fitness in a theoretically useful sense (Abrams, 2012; Beatty & Finsen, 1989; Brandon, 1978, 1990; Mills & Beatty, 1979; Pence & Ramsey, 2013; Rosenberg, 1982; Sober, 1984, 2001; Wagner, 2010). If there is nothing more to the notion of fitness than realized fitness, then the claim that "fitter organisms reproduce in greater number than their less-fit conspecifics" is equivalent to "the organisms that reproduce more out-reproduce those which

✉ Peter Takacs
peter.takacs@sydney.edu.au

Pierrick Bourrat
p.bourrat@gmail.com

[1] Department of Philosophy and Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia

[2] Department of Philosophy, Macquarie University, Sydney, New South Wales, Australia

reproduce less." As stated, this depicts natural selection as a tautology that is devoid of empirical content and, thus, explanatory import. It fails to account for why any of the organisms reproduce in greater or lesser numbers than their conspecific competitors. Even more worrisome for the naïve view is that it routinely generates evolutionary scenarios in which the fitnesses of organisms exhibiting the same selectively relevant trait type are inconsistent within a homogeneous selective environment (Brandon, 1990; Scriven, 1959), a consequence which threatens to undermine the generally agreed upon role of fitness as a causal property in selective explanations of adaptive evolution.[1]

The propensity interpretation of biological fitness (hereafter "PIF"), originally championed by Robert Brandon (1978) and Susan Mills and John Beatty (Mills & Beatty, 1979), presented a way to maintain a static notion of organismal fitness without the drawbacks inherent to the naïve view. Brandon and Beatty (1984) summarize the upshot of this maneuver:

> On the propensity interpretation, fitness is a probabilistic disposition or ability explicated in terms of expected rather than actual reproductive success (in the mathematical sense of 'expected value'). Inasmuch as the connections between an entity's dispositions or abilities and its actual behaviors are causal connections rather than analytic connections, the propensity interpretation of 'fitness' allows for genuinely explanatory accounts of differential reproduction in terms of differential fitness […] Informally, 'fitness' is defined in terms of abilities to reproduce, not in terms of actual reproductive success. (pp.33-34)

Their proposal sidesteps the explanatory worries that accompany the naïve view's commitment to defining fitness simply in terms of realized fitness. Unlike the naïve view, it does not conflate evidence for the presence or value of a probabilistic property with the property itself.[2] Moreover, it acknowledges the formal methods (e.g., multivariate linear regression) that biologists routinely deploy to ascertain fitness values.

Early formalizations of a static view of fitness nonetheless foundered on problems associated with demographic and environmental stochasticity (Gillespie, 1977). Selective scenarios involving within-generation and between-generation variance in offspring number featured prominently in philosophical work (Beatty & Finsen, 1989; Brandon, 1990; Sober, 2001). Fitness assessed solely on the basis of the arithmetic mean number of offspring overlooks the effects of variance (as well as potentially important effects of higher mathematical moments such as skew, kurtosis, etc.) for explaining and predicting evolutionary dynamics.

---

[1] The claim that organismal fitness is a causal property can, of course, be denied. While we cannot here address the nuances of this "statisticalist" view, the challenges posed deserve serious scrutiny (cf. Otsuka, 2016; Walsh et al., 2017).

[2] See Beatty and Finsen (1989) for more about the so-called "operationalist fallacy."

It was in the face of such seemingly insuperable problems that Ramsey (2006) reiterated the need for a static conception of fitness for selective explanation.[3] He labeled this invariant notion "Block Fitness." Pence and Ramsey (2013) later provided a rigorous explication of the concept. A major obstacle to the acceptance of Block Fitness even in its more mathematically rigorous incarnation is that practicing population biologists more often use dynamic fitness functions, which are prima facie incompatible with or uninformed by it. Situations involving negative frequency-dependent selection, an example we will return to later, present a simple illustration. In such situations, the fitness of an individual decreases as the frequency of its type increases in the population (Ridley, 1996, pp. 127–129).

Abrams (2009a, 2009b, 2013, 2014) has recognized this tension and instead argues for a *conditionalized* static notion that purportedly reconciles the apparently dynamic fitnesses used in biological practice with the explanatory invariance required by a theory of adaptive evolution via natural selection. In this paper, we show that, despite the initial tenability of Abrams's proposal, it can lead to conflicting intuitions about whether selection actually occurs in a population. While we ultimately find Abrams's account wanting, it is not our aim to relieve the existing tension in favor of Block Fitness and PIF. Nor is it our intention to suggest that the tension is irresolvable. Instead, we seek only to (i) highlight the theoretical commitments and shortcomings that come with each and, thereby, (ii) raise the possibility that there may be an unavoidable epistemic trade-off when opting for one static conception of fitness rather than the other.

2 Block fitness formalized

The prospect of finding an adequate formalization for a static conception of fitness were bleak until Pence and Ramsey (2013) introduced a new mathematical model to underpin the philosophical interpretation provided by Ramsey (2006). Ramsey contends that "Block Fitness—is a block property, constant over an individual's lifetime, implying that no events in an individual's life, altruistic or otherwise, can change its fitness. Fitness, then, is not a function of the actual life that an organism lives. Rather, […] it is a function of all of the possible lives that the organism could have lived" (2006, p.485). Drawing on results in a research program known as "adaptive dynamics" (Metz et al., 1992; Tuljapurkar, 2013), Pence and Ramsey offer a general model to formally explicate this notion of fitness.[4] For our purposes, it is of the utmost importance to note that their model is intended as a definitional measure of *individual* fitness (Pence and Ramsey's '*F*'), as in the fitness of a token organism. As with earlier versions of the PIF, an individual's fitness is still to be established by applying probability weightings (from census data) to a distribution of

---

[3] For convergent thoughts, see Bourrat (2015, 2017) who proposes that natural selection (as opposed to other evolutionary processes) results from differences in intrinsic-invariable properties (i.e., static properties) that lead to differences in reproductive output.

[4] For details, see Equation 4 on p.862 of Pence and Ramsey (2013).

outcomes. But the outcome states under consideration are no longer confined to the possible (whole numbered) reproductive contributions of a parental generation to a subsequent generation. The outcomes instead consist of all the "daughter populations" (or lineages) to which an organism with genome, G, in environment, E, might give rise.[5] The guiding idea is more straightforward than this wording suggests. Take any organism in a population. There are many different ways that an organism's life could unfold. It might, for example, give rise to many progeny or it could leave relatively few. It could just as easily leave none at all due to its succumbing to severe malnutrition or predation prior to reproduction. There are also many ways that a particular organism could produce one and the same number of offspring. For instance, it might reproduce at an earlier rather than later time in its life. The same goes for any of the offspring or even grand-offspring of our focal organism. Radical contingency of this sort holds for any possible descendants. So, the total number of "daughter populations" or "possible lives" is, strictly speaking, uncountable. And the set containing all such daughter populations (Pence and Ramsey's "Ω") is the total state (sample) space.

Let us restrict focus to the total sample space containing all possible daughter populations, including the actual daughter population (or lineage) to which an organism gives rise. This space is explicitly designed to cover all possible sets of circumstances within which a focal organism and its descendants may find themselves. The selective environment, on this propensity interpretation, consists of all possible background conditions that are potentially sampled by competing trait variants or individuals in the long run. Some possibilities (or regions of state space) can become more probable than others once information about the individual, the population(s) to which it belongs, and the environment(s) it inhabits is taken into account. The idiosyncrasies that distinguish one selective environment or scenario from another in effect generate "special cases" of this more general measure. More crucial for our purposes, though, is that the fitness proposed by Pence and Ramsey is static since its true measure can only be determined in the very long run. What they offer, then, is nothing short of a "God's-eye-view" of individual fitness.[6]

It is worth dwelling on this point for a moment. Perhaps the most basic explanatory relationship within evolutionary biology is that of "being better adapted than" or "fitter than" within a particular population (Bouchard & Rosenberg, 2004). All systems subject to the evolutionary process are arguably governed by the common expectation that if organism X is better adapted than organism Y in environment E, then (probably) X will have more (sufficiently similar) offspring than Y in E (Brandon, 1978). Brandon (1990) calls this basic assumption "the principle of natural selection" on the grounds that it must be presupposed when explaining the dynamics of any evolutionary system. If we wish to maintain that natural selection is an objective

---

[5] Op. 860-861 of Pence and Ramsey (2013). What follows draws very heavily on Pence and Ramsey's wording.

[6] The adjective "God's-eye-view" is not intended as pejorative. It is illustrative rather than evaluative. There is nothing philosophically untoward about attempting to establish metaphysical foundations for evolutionary theorizing.

feature of the natural world, one largely if not exclusively responsible for adaptive evolution, we require an objective conception of fitness. Pence and Ramsey's static conception provides a cogent, even if contentious, metaphysical interpretation of fitness as an unchanging objective property of token individuals. It grounds fitness as the causal fodder for selection and thus moves us beyond the unsatisfactory assumption that fitness is nothing more than a theoretical primitive (Rosenberg, 1983; Rosenberg & Williams, 1986; Williams, 1970).

3 Counterintuitive consequences of block fitness and a conditional static alternative

As convincing as Block Fitness may be from a philosophical standpoint, its detractors are quick to point out that its appeal is perplexing considering the actual practice of biologists. A litany of studies in evolutionary population biology reveals that fitness values are often taken to be variable or dynamic features of trait types or token organisms.[7] The phenomenon of frequency-dependent selection presents a case in point. Frequency-dependent selection occurs whenever the fitness of a trait changes as a result of changes to its frequency in the population.[8] Density-dependent selection, wherein selection among competing trait types in a population is affected by how many individuals of the species occupy the habitat, is just as crucial to explaining and predicting evolutionary dynamics. Therein, too, relative fitness values are often portrayed as dynamic rather than fixed. And it has long been known how unpredictable fluctuations in the biotic or abiotic components of the environment (e.g., weather events) can also affect the ordering and magnitude of relative fitness differences in a population (Levins, 1968).[9]

For illustrative purposes, let us examine the phenomenon of *negative* frequency-dependent selection, which occurs when becoming more common makes a trait, along with the organisms that bear it, less fit. Examples of negative frequency-dependent selection abound. Within biological game theory, which focuses exclusively on phenotypic traits (or strategies) and the optimal outcomes (or equilibrium payoff states) of episodic natural selection, an organism playing an aggressive or confrontational strategy known as "hawk" will flourish in a population composed primarily of passive players known as "doves." However, a dove will tend to do better in a population composed mainly of hawks (Maynard Smith, 1982). It has recently

---

[7] Although studies involving trait types feature prominently, it is worth remembering that this contingent fact is a matter of "occupational necessity" designed to prevent the conflation of individual fitness (as a propensity or propensity-like property) with realized individual fitness. It can be argued that biologists aim to determine the fitness of token organisms, which can only be approximated by way of indirect inferences that rely on selectively relevant trait types. On this point, compare Sober (2013) and Pence and Ramsey (2015).

[8] It is worth noting that models for the evolution of altruism rely on a form of frequency-dependent selection.

[9] Pence and Ramsey are clearly aware of the difficulties raised in this section. They explicitly state, "Our Equation (4) is the density-independent, non-chaotic limit of this more sophisticated work [in adaptive dynamics]" (2013, p.863).

been shown (Kokko et al., 2014) that the head-color polymorphism of red and black Gouldian finches (*Erythrura gouldiae*) strongly supports the game-theoretical model for limited aggression in an animal population. The aggressive red morph is behaviorally dominant and successfully invades black populations, but when red "hawks" become too common, their success is severely compromised via decreased parental ability. Other studies have demonstrated negative frequency-dependent selection for self-sterility alleles in plants, handedness in the feeding morphology for scale-eating cichlid fish, and male reproductive strategies in a marine isopod (Stevens, 2011).

Despite Pence and Ramsey's responses to possible objections concerning the relationship between short-term and long-term fitness, an outstanding problem for their infinite-timescale measure (2013, p.868) is that it is unlikely to correspond to the relative fitness measures that drive selective dynamics in the ecological and evolutionary short term (~1-100 generations). Population biologists are often interested in how critical demographic rates (birth rates, death rates, life expectancies, age of reproductive maturity, etc.) or the relative frequencies of competing trait types respond to inter- as well as *intra*generational changes in intrinsic (e.g., intraspecific density) and extrinsic (e.g., resource availability or predator abundance) factors. Static fitness values calculated over an infinite amount of time and an uncountable sample space seem of less value to such research, as is evinced by researchers' readiness to countenance a dynamic notion of fitness. Moreover, it is not at all obvious, as Ramsey (2006) contends, that *any* measure of realized fitness, whether for a token individual or as the expected fitness of an ensemble (i.e., type fitness), "is not very predictive" (p.489). In cases of negative frequency-dependent selection, as with the Gouldian finches noted above, selective dynamics can be accurately predicted within as well as across generations in the short term by deriving expected fitness exclusively on the basis of realized values. While such a measure might not render accurate predictions over longer periods, it could nevertheless suffice as a bona fide short-term predictor (Ariew & Ernst, 2009; Ariew & Lewontin, 2004; Walsh, 2007, 2010).

Largely in response to the foregoing problem, Abrams (2009a, 2009b, 2013, 2014) has proposed a way to alleviate the tension between the static and dynamic conceptions of fitness. Although critical of propensities (Abrams, 2007), he recognizes the importance of upholding a static conception of fitness if we are to maintain that natural selection is an objective feature of the natural world that forges organismal adaptation. To resolve this tension, Abrams argues that "the relevant probabilities [with respect to fitness] are over 'organism–environment histories': sequences of states of an organism and its environment. [His] approach explicitly concerns organism types rather than tokens" (2009a, pp. 489–490). Organism-environment histories are underwritten by what Abrams calls "static conditional fitnesses." These are measures of relative fitness "conditional on the occurrence of an event of a particular type. Though these fitnesses do not change during an organism's life, they reflect the effects of possible events on organisms' reproductive success" (2009a, p. 490). In other words, following Abrams, tokens of one and the same organismal type (i.e., individuals who exhibit the same character state or trait variant) can have different fitness values only if there are corresponding nonequivalent descriptions of the conditioning environment.

An example inspired by Abrams (2009a) is instructive. Imagine that individuals of a certain species are more likely to survive to sexual maturity (S) if they happen upon a rare nutrient-rich resource (E), such as carotenoids, in their habitat.[10] All else being equal, S should track fitness. Suppose that there are just two discrete heritable trait variants (A and B) in our hypothetical population. While both trait variants tend to do better when encountering carotenoids, it turns out that individuals with variant A gain an even greater (fitness) advantage than do individuals with variant B. However, in the absence of carotenoids, individuals with variant B have a significant advantage (i.e., incur a much smaller cost to fitness) over individuals with variant A. Let us further assume that there is no correlation between exhibiting a particular trait type and whether an individual will encounter such a resource. At first glance, it appears as though the viability fitness (S) of an individual can change on the basis of a fortunate encounter with E. This goes for individuals of both trait variants. But this would be inconsistent with a static notion of fitness. Abrams resolves this apparent inconsistency by arguing that if an organism's survival is conditionalized on the relevant occurrence (E) or nonoccurrence (~E) of encountering the nutrient rich resource, it can still refer to a static property:

(I) $P(S_A|E) > P(S_B|E)$
(II) $P(S_A|\sim E) < P(S_B|\sim E)$

It is unsurprising for relative fitness differences, expressed here as probabilistic inequalities, to change in magnitude or direction *when the background conditions in which natural selection occurs are nonuniform.*[11] The inequality in (I) shows that individuals with trait A have a greater relative fitness when a nutrient rich resource is encountered, and vice versa for individuals with trait B in complementary circumstances (II). Some A-type individuals in E will undoubtedly perish despite finding a nutrient-rich resource. But their fates are supposedly accounted for by a statistical measure of variance in fitness, which is, in turn, the average difference between observed and expected fitness conditional on E. Abrams makes the overall point unmistakably clear: "[T]he relevant probabilities […] are for a given organism type in a given environment, and these probabilities don't change within organisms' lives, since they already take into account everything that can happen in a life, i.e., in various possible lives" (p.493).

4 Is Conditionalized static fitness the best of both worlds?

Abrams presents what might be considered a static Humean (complex type) alternative to Pence and Ramsey's Block (token) fitness. His organism-environment histories can be construed as composites (combinatorial or topological possibilities)

---

[10] Astute readers will recognize that this is a slightly modified version of Abrams's (2009a) example involving carotenoids.

[11] See Brandon (pp.60-64, 1990) for details about, what he calls, the "selective environment."

consisting of (1) "evolutionary individuals" who are type-cast as mereological fusions via reference to selectively relevant, determinable traits and (2) types of extrinsic events with corresponding complementary or contrast classes. Explaining evolutionary dynamics in any population thereby requires reference to probabilistic regularities (e.g., net relative fitness differences) based on the types that are used to describe an organism-environment history. The explicit aim is to ground the invariance required for an explanatorily useful notion of fitness and to do so in a way that mirrors the practice of population biologists who routinely deploy dynamic measures. While containing much to admire, we will argue in this section that Abrams's proposal suffices only if one is willing to pay, what is by our accounts, an unacceptably high epistemic cost. More specifically, as we will show, his proposed resolution privileges a form of explanatory pragmatism that compromises the attainment of a determinate biological ontology. We now introduce an example that is formally identical to the previous example and highlights the tensions to which we have thus far alluded.

Upon observing distinct character states (trait variants), an evolutionary biologist may want to understand how or why such variation is maintained. Is it just a case of "neutrality" in the sense that there are no intrinsic or extrinsic selective pressures associated with this trait? Perhaps the observed variation is due to epistatic,[12] pleiotropic,[13] or developmental constraints on the efficacy of selection. Or is it the result of countervailing extrinsic or intrinsic factors (e.g., a mutation-selection balance)? At any rate, testing competing hypotheses such as these is commonplace in evolutionary population biology. In this hypothetical scenario (Table 1 generated from data in Appendix), there are two mutually exclusive and exhaustive heritable trait variants: A and B. The continuous variable "Available Nutrients" has been arbitrarily discretized by establishing an inclusive threshold value for "Nutrient Rich", which generates the complementary contrast class consisting of any (nonnegative) value below the threshold. The two resulting environmental conditions (E1 and ~ E1) are equiprobable and randomly inhabited, so there is no correlation between trait type and environmental condition. In other words, embodying a particular trait variant has no effect on the probability of encountering either type of nutritional environment. Fitness coefficients (w_A and w_B) here designate *relative* fitness, a standard measure of the (dis)advantage of one variant against its extant competitor(s). By convention, the trait variant with the highest relative fitness in each selective environment becomes a reference trait for normalization and is ascribed a value of 1.0. The strength of selection against the less fit trait variant (i.e., selection coefficient) can, then, be easily deduced as departure from 1.0. Arithmetic means (μ_i) and variances ($\sigma^2{}_i$) in reproductive output are provided to make evident that higher moments of the fitness distribution "cancel out"[14] when assessing relative fitness and selective

---

[12] Epistasis occurs when alleles at two or more genetic loci interact non-additively to determine the phenotype.

[13] Pleiotropy occurs when one genetic locus affects more than one phenotypic trait, which causes a genetic correlation.

[14] In other words, the fitness distributions for the character states are identical.

dynamics for a more coarse-grained (i.e., global) description of the population in which $w_A = w_B$. Contrary to what some might expect, this example also shows that global variance(s) can increase despite a shift to larger population size (see Table 1). The values used in this example expressly circumvent many of the difficulties that Pence and Ramsey's (2013) model for individual fitness was designed to resolve. Their model would accordingly be much simplified when applied to a scenario like ours.

There is an obvious fitness reversal for a partition of the population by one environmental condition to the exclusion of the other:

(III) $\Pr(w_A | \mathrm{E1}) > \Pr(w_B | \mathrm{E1})$
(IV) $\Pr(w_A | {\sim}\mathrm{E1}) < \Pr(w_B | {\sim}\mathrm{E1})$

Trait type A is fitter than trait type B in nutrient-rich conditions and vice versa in nutrient-poor conditions. Furthermore, the magnitude of the relative fitness advantage that accrues to A in E1 is precisely equal to that which accrues to B in ~E1. While this is undoubtedly a somewhat "contrived case" that follows from stipulated features of our example, similar instances could be generated even when the magnitudes of the relative fitness differences are unequal. Now, a static measure of fitness cannot accommodate such a blatant inequality reversal if it is to figure as a causally operative factor in a selective explanation (Walsh, 2007, 2010). Note, moreover, that the fitness reversal in this case cannot be brushed aside as the result of what some have deemed "illegitimate averaging" (Glymour, 1999; Otsuka et al., 2011). Partitioning the population by way of E1 and ~ E1 does not yield partitions of unequal size, nor does it introduce any type of confounding dependency between nutrient level and trait variant. The number of individuals assigned to each partition is the same, and the number of A-type variants is stipulated as matching the number of B-type variants within each partition. Individuals with variant B are no more or less likely to encounter a particular type of nutrient level (E1 vs. ~E1) than conspecifics with variant A. Following Abrams's proposal, then, the fine-grained partition as in (III) and (IV) demonstrates that fitness can remain static if conditional on homogeneous selective environments E1 and ~ E1.

We of course agree with Abrams that fitness must be assessed in homogeneous selective environments. A pressing difficulty remains, however. The problem is that the selective environment is homogeneous at both the local level and the global level in our example. As both ways of partitioning the setting refer to the same individuals, these individuals supposedly undergo the same selective pressure(s) no matter how we might choose to describe the setting. If homogeneity is the sole criterion for determining the "correct" conditionalization, we are faced with the dilemma of having to choose between two legitimate ways of conditionalizing to achieve homogeneity. The difference is anything but trivial. Opting for a fine-grained partitioning as on the local level,[15] a practicing population biologist must conclude that natural

---

[15] This local interpretation or "fine-grained partitioning" is equivalent to situation in which there is a metapopulation consisting of two subpopulations of equal size and there is no migration or mutation.

Table 1. A Scenario for Selective Incompatibility (See text and Appendix for details.)

Interpretation of Selective Environment: GLOBAL	**Interpretation of Selective Environments: LOCAL**
ENVIRONMENT (E) (Available Nutrients) $w_A = w_B$ (1.0 = 1.0) ($\mu_A = \mu_B = 8$; $\sigma^2_A = \sigma^2_B = 4.8$)	**NUTRIENT RICH (E1)** $w_A > w_B$ (1.0 > 0.6) ($\mu_A = 10$; $\mu_B = 6$; $\sigma^2_A = \sigma^2_B = 0.8$) **NUTRIENT POOR (~E1)** $w_A < w_B$ (0.6 < 1.0) ($\mu_A = 6$; $\mu_B = 10$; $\sigma^2_A = \sigma^2_B = 0.8$)

selection occurs because A is fitter in E1 and B in ~E1. This might, for example, indicate that there is a trade-off (i.e., a negative phenotypic correlation). On the global level, in contrast, a practicing biologist would posit E (nutrient availability) as part of the background rather than as an explicit variable upon which to condition, which implies that all individuals are subjected to the same conditions in equal measure. The reversal in fitness being of the same magnitude on the local level, however, this biologist would be forced to conclude that there is no evidence of natural selection. A and B are selectively neutral from the global standpoint; character state apparently makes no difference to fitness. There is, consequently, unresolved arbitrariness surrounding how to answer the question "does natural selection occur in this population?" This indecision generates inconsistent predictions about what would happen if available nutrient levels were to fall predominantly above or below the discrete threshold identified in our example.

Resorting to the fiat of researchers, without additional qualification, cannot resolve this worry. We can begin to see why by noting the variance around mean reproductive output after partitioning by E1 and ~ E1: $\sigma^2_A = \sigma^2_B = 0.8$. There is variance on the global level ($\sigma^2_A = \sigma^2_B = 4.8$) as well. This indicates that there are individuals within each trait type whose actual fitness deviates from the expectation in either partitioning. How should we accommodate individuals exhibiting, for example, trait type A in E1 with unequal realized fitness? Following Abrams,[16] there must in principle be yet another environmental variable by which to partition. Let us call it "F1" and its complement "~F1." We would, then, be faced with an even more fine-grained partitioning, one in which relative fitness differences must be assessed in four different homogeneous environments: E1and F1; E1and ~ F1; ~E1 and F1; ~E1 and ~ F1. Now, it is plausible if not likely that variance in reproductive output will remain even after such a refinement. If so, yet another round of partitioning would accordingly ensue. The process of revision via complex partitioning presents a slippery slope; it could be repeated indefinitely because the token causal events that

---

[16] We do not mean to imply that Abrams is oblivious to this problem. See, especially, Abrams (2009b) for a suggested resolution to, what he calls, "the problem of the reference environment."

influence the prospects for survival and reproduction in any actual environment is effectively infinite.

This regress can terminate in one of two ways. It must either proceed until discovering a partition that yields no variance in fitness for either A or B, in which case it becomes a deterministic setting, or stop short of this limit for reasons associated with the inclinations and biases of knowledgeable researchers (e.g., returns on investment of time or funding, progress towards answering questions addressed, etc.) because, perhaps, the setting is fundamentally indeterministic.

To his credit, Abrams (2007, 2009b, 2014) recognizes that an unchecked regress is unpromising. Considering the number of variables involved, a completely exhaustive partitioning of this ilk would recast each individual in a population as experiencing a completely unique environment, which then renders an explanation of relative success (i.e., differential fitness) useless. Abrams refers to such extremely specific, idiosyncratic environments as "microenvironments" or particular "circumstances" (2007, 2014) in order to distinguish them from recurring sub-environments. For him, only the probabilities associated with the occurrence of the latter are relevant to explaining selective dynamics.[17] Recurring sub-environments would indeed prove more explanatorily relevant, but the pressing issue here is one of whether we can for principled reasons maintain consistent static fitness values in partitions considerably less specific than "microenvironments." Abrams apparently thinks that we can:

> [F]or a subenvironment to be relevant to natural selection in a small population, it must encompass a broader range of circumstances, *ceteris paribus*. These relationships could be formalized in terms of a particular model, producing a formula allowing calculation of probabilities of recurrence necessary to produce a noticeable effect of selection given values for the parameters just mentioned […] For a given population and environment, and a specified level of likely effect of selection, such a formula could in principle be used to estimate a *minimal environmental grain*: a specification of how fine-grained a partition of the environment can be while still making subenvironment-relative fitness differences themselves relevant to selection.
> I'm not sure how useful such an estimation project would be, however. The main point is that modeling practices and empirical research are consistent with a vague boundary between subenvironments that have a significant probability of recurrence, and those that don't. Above this vague limit, researchers are free to choose a way of partitioning environmental conditions into subenvironments that is useful for their research goals. (2014, p.19)

The basic idea is that biologists can partition in whatever manner they see fit so long as they do not do venture "below" the limit at which the selective environment becomes unshared (i.e., heterogeneous). But establishing the possibility of a

---

[17] Op.20 (Abrams, 2014): "What matters to natural selection is not this or that organism's particular circumstances and particular fate, but the sorts of conditions that individuals in the population are likely to encounter repeatedly."

minimal environmental grain in the manner Abrams suggests ignores the difficulty that our example (Table 1) is designed to highlight. We have restricted ourselves to identifying just two of possibly many distinct ways that biologists might partition the selective environment *above* Abrams's "vague limit." Our example raises a problem for any case in which the *minimal* environmental grain is not also the *maximal* environmental grain of description. If (i) researchers are "free to choose a way of partitioning environmental conditions," (ii) the various descriptions they deploy supposedly correspond only to distinct *aspects* of a single mind-independent reality, and (iii) selection is taken to be an objective feature of the world, then there is no room in the end for inconsistent answers to the question "does natural selection occur?" when posed of any actual population. Attempting to resist this by arguing that the question admits of inconsistent answers under distinct descriptions simply elicits the following rejoinder: "which description is best for the purpose of explaining and predicting evolutionary dynamics *in this (actual) population*?" Returning to our example, this question is tantamount to asking whether to include "nutrient level" (i.e., E1 or ~ E1) as an explicit variable in a model of selective dynamics. If it is included, countervailing selection occurs. If it is excluded, there is no selection. Leaving this issue to the vagaries of "modeling practices and empirical research" fails to offer a principled resolution to this "promiscuity of (acceptable) partitions" problem. Stopping the regress (short of Abrams's microenvironments) apparently entails an unbearable excess.[18]

For Abrams, determining a proper partition and thus retaining a static value of fitness seemingly depends on the biases and inclinations of researchers—that is, "research goals and available data and methods" (2013, pp.297-298). Opting in this direction seems to accord quite well with what often transpires as population biologists come to better understand the systems confronting them. Unfortunately, we then again find ourselves in a situation that is not relevantly different from the one our example was designed to highlight: two knowledgeable evolutionary biologists disagree about the level of proper partitioning (i.e., homogeneous selective environment). And, as before, the mere fiat of researchers seems an unsatisfactory desideratum for determining whether fitness is an objective causal property and selection an effectual process.

Now, we would be remiss if we failed to mention that Abrams (2013) has addressed a less detailed but formally similar case to the one we present in Table 1.[19] He there considers how the probabilities associated with randomly assigning birds (English magpie pigeons) with competing color patterns to mutually exclusive environmental conditions (dark sand vs. light sand) can figure in selective explanation.

---

[18] It also threatens to blur the distinction between drift and natural selection, an outcome that runs counter to the intuitions of most biologists and philosophers (Brandon, 2005). If your mathematical expectations differ substantially, then so, too, will your views about which outcomes are due to selection (i.e., are adaptions) and drift (i.e., departure from expectation).

[19] "*Second scenario*: Imagine a single environment in which sand color varies in patches whose average diameter is about a meter. The random assignment of pigeons to a background sand color is then determined by where pigeons happen to land. In either case, we suppose that the pigeons are not selective about sand color" (Abrams, 2013, p.295-296).

His proposed resolution makes explicit two important features that can be juxtaposed with our example. How and why we disagree with Abrams's can be further clarified by examining these.

The first feature is an unobjectionable epistemic claim about how biologists can come to understand or investigate selective dynamics. Abrams urges that evolutionary biologists should be granted considerable autonomy when describing the environment(s) inhabited by a population. He nevertheless recognizes that there must be some constraints. For instance, biologists should avoid descriptions that introduce confounding factors that lend to instances of "illegitimate averaging." They should, as per Abrams's suggestion, restrict themselves to homogenous partitions of the environment at or above the minimal environmental grain of description. Our example (Table 1) accordingly partitions the population only via statistically "legitimate" descriptions of the environment with a nonnegligible probability of occurrence. Investigating the ramifications of alternative descriptions can, as we have already noted, reveal the presence of countervailing causal influences (e.g., life history trade-offs). This commitment to what might be called "bounded exploratory liberty" is shared by many, some of whom part ways with Abrams's conclusions regarding the metaphysical status of fitness (Ramsey, 2006; Walsh, 2010). We likewise concede this general epistemic point to Abrams. Members of the biological community should retain the right to (defeasibly) *depict* biological entities and environmental factors however they see fit. To prematurely withhold this liberty is anathema to prospects of theoretical progress.

The second feature of Abrams's position proves a different matter. Abrams contends that descriptions at any level can single out and, thereby, emphasize distinct causal relations.[20] This claim is prima facie unproblematic. Close examination of any actualized population will reveal a complex causal nexus (i.e., a highly interactive network of causal relations across many levels of organization). Of course, only some these causal relations will be relevant for explaining the evolutionary trajectory of a population. Even within the set of causal relations that are deemed explanatorily important, there may be pragmatic reasons (e.g., matters of research interest) for emphasizing some causal relations at the expense of others. These differences can be reflected in the use of nonequivalent conditioning descriptions of the sort that Abrams condones, so long as they arise above his vague limit for the minimal environmental grain. Notice, however, that the case presented in Table 1 includes two legitimate descriptions that refer to the same causal substrate. These descriptions lead to competing, incompatible evolutionary explanations. Our contention is that the phenotype exhibited by one and the same individual in any actual population

---

[20] "I argue for the stronger claim that it's plausible that: [1] Different descriptions of a biological population sometimes pick out distinct effects, and [2] There are distinct causes producing these effects. This is a claim about causation rather than explanation. Causal relations exist independently of how the population is characterized, though our way of describing the population can, indeed, focus on one or another of these causal relations" (Abrams, 2013, p.295). Later in the same paper, he claims that: "My view implies that a researcher's description is able to pick out different causal relationships, but allows that these causal relationships exist, realized by some the of same parts of the world, independent of the choice of causal relationships on which to focus" (p. 300).

ought to be either adaptive or not. Whether it is should not depend on our preferred description of the conditioning environment. We consequently find Abrams's suggested resolution wanting.

Abrams can respond to this in one of two ways. He can either (i) hold that the partitioning descriptions (global vs. local) applied to an actual population do not in fact distinguish distinct evolutionary explanations, or (ii) opt for one level of description over the other.

Going the way of (i) fails to address the case presented in Table 1. Our case is designed expressly to contrast two competing evolutionary hypotheses, ones which obviously make inconsistent predictions about what would happen if nutrient level shifts above or below a specified threshold. So long as our investigation is confined to the evolutionary dynamics of a single actualized population, one of these descriptions must be rejected.

The only sensible alternative for Abrams is option (ii). What he calls an "exclusive description" (2013, p.295) would correspond to the subpopulation picked out by the fine-grained descriptor "E1" in our example. This exclusive description can reveal the population dynamics that would ensue if there was no chance (or only a negligible chance) for individuals to encounter ~E1.[21] While supposedly depicting *causal relations* between A-type or B-type individuals and a particular environmental condition E1, Abrams asserts that an exclusive description nevertheless falls short of being a relevant evolutionary *explanation*.[22] This follows because an exclusive description identifies the static fitness effects of trait variants in what is an incompletely described selective environment. Uncomplemented by ~E1 in our example, for instance, E1 would not be the selective environment shared by all members of the population. It is, at best, a hypothetical (counterfactual) selective scenario that speaks to what the fitness effect of a trait type would be if it occurred in the sort of selective environment identified by the exclusive description. According to Abrams, providing an adequate explanation of selective dynamics in our population requires an "inclusive description" (2013, p.295). The inclusive description captures the causal relations in *each and every one* of the local (non negligibly recurring) partitions allowed by the environment. An inclusive description is thus the union of the exclusive partitions that are available on the local ("fine-grained") description. In our example, the inclusive description must include E1 as well as its complement ~E1. Abrams contends that this information is crucial for explaining and predicting population dynamics because evolution is a spatially as well as temporally extended process. If the individuals exhibiting a trait type have a nonnegligible chance of inhabiting exclusive but allowable partitions of the environment,

---

[21] In an earlier paper, Abrams says the following: "Rather than an organism's fitness literally changing when an effect C occurs, there is instead a relationship between conditional probabilities: a type's fitness conditional on the occurrence of C is different from its fitness conditional on C's non-occurrence" (2009a, p.493).

[22] On this point, Abrams (2013, p.296, emphasis added) says the following: "It should be clear, though, that if what we are concerned with is the evolution of the pigeon population over several generations […] then *any relevant sense of fitness and selection must take into account probabilities of pigeons being found on light or dark sand when hawks are overhead*."

we must accordingly determine just how "accessible" these distinct environmental conditions are. In other words, we need to know the probability of encountering a specific environmental partition given that an individual exhibits a particular character state.

If Abrams opts in favor of the inclusive description (i.e., the "local interpretation" with two environments from Table 1) in our case, his view encounters other difficulties. The design of our example makes it such that there is a nonnegligible chance of 0.5 that individuals of either type will encounter E1 and ~ E1. Half of the A-type individuals will find themselves in E1, the rest inhabit ~E1. The same goes for B-type individuals. The presumption is that this "random occupancy" will continue into the future. But the probabilities associated with a character state's encountering E1 and ~ E1 in our case could also have been generated temporally (e.g., from longitudinal data on transgenerational occupancy) rather than spatially (e.g., from cross-sectional data on within-generation occupancy). In such a temporally extended scenario, A-type (B-type) individuals might have been restricted to randomly occupying only one type of sub-environment per generation. That is, A-type (B-type) individuals could occupy either E1 or ~ E1 (but not both) during any particular generation. Since individuals adopt the statistical properties assigned to their character state type, an individual currently residing in E1 would then still have a nonnegligible probability (= 0.5) of being found in ~E1 *without it being the case that it is actually in ~ E1*. In that case, the static fitness values derived from the exclusive partition E1 (or alternatively ~E1) would provide a better explanation and more informative predictor of selective dynamics in the short term (Doulcier et al., 2021). Insofar as researchers' interests dictate not only the locations but also the durations over which selective dynamics are to be assessed, they also govern what information from allowable partitions is explanatorily relevant. A description might be considered "inclusive" by students of microevolution at a timescale of 100 generations even though it tracks selective dynamics exclusively in E1, but "non-inclusive" (i.e., "exclusive") by students of macroevolution when it ranges over 10,000 generations because the probability of finding a trait type in ~E1 is no longer negligible over this latter (macroevolutionary) timescale. Contra Abrams, then, it seems as though the information from an uncomplemented local partition (i.e., an "exclusive description") can suffice as a relevant explanation of selective dynamics for biologists (e.g., conservation biologists) who wish to make predictions about what will occur only in the short term. Further, decisions regarding the applicability of an exclusive description (i.e., whether the set of causal relations in a local partition is explanatorily relevant) seem to presuppose epistemic access to an inclusive description. An exclusive description is, as per Abrams's suggestion, explanatorily relevant only if its (nonnegligible) complement is made explicit. This runs counter to the situation confronting many if not most practicing biologists. They often have no such access. Alternatively, they can decide that such access, even if attainable, is unnecessary in light of their explanatory interests.

In response to our claim that inclusive descriptions do not map neatly onto the ways that biologists use fitness, Abrams could always maintain that the philosophical project of explanation is distinct from the goal of finding measures that ground accurate prediction (Millstein, 2016). However, here are drawbacks to this approach.

For starters, recall that what initially sets apart Abrams's static account of fitness from Pence and Ramsey's "Block" alternative is that it supposedly captures the actual practice of biologists who routinely countenance dynamic fitness measures. His account unarguably loses some of its contrastive luster if this is sacrificed.

An even more daunting concern, though, is whether Abrams's account can avoid collapsing into the "Block" account by promoting the use of inclusive descriptions. For the reasons discussed in the paragraph before last, there can be considerable disagreement among evolutionary population biologists about what constitutes a properly "inclusive" description. Ever more inclusive descriptions seem readily available. The more fine-grained an inclusive description becomes, however, the more it will "fill in" (make explicit) the details that a global description presumes as background conditions. Our example shows how this can occur. The union of complementary partitions in Table 1, which yields an "inclusive" description by Abrams's standards, collapses the explanatorily relevant causal relations into those that are captured by the global level interpretation of the selective environment.

5 Conclusion

We have argued that a tension remains between the static and dynamic conceptions of fitness. Any theoretically informative notion of fitness is unavoidably conditional. The issue of importance is how to determine the proper partitioning for conditionalization. If fitness is explicitly conditionalized on some but not all environmental circumstances, as with either E1 or ~ E1 in our example (Table 1), then it is static only over some *part* of a complete probability distribution. While Abrams's account may thus accord well with the actual use of (relative) fitness as a predictive and explanatory apparatus, it nevertheless fails to provide principled means by which to determine whether or to what extent natural selection occurs when there are incompatible homogeneous selective environments (i.e., E "global" as opposed to E1 and ~ E1 "local"). His proposal consequently suffers from an intolerable form of descriptive arbitrariness that threatens explanatory inconsistency.

Alternatively, if fitness is conditioned on all the possible circumstances that can affect survival or reproduction in the infinite limit, it is thereby conditioned on Pence and Ramsey's sample space (Ω), which includes every possible selective scenario. In our example, this would be tantamount to stipulating that the local level partitioning (i.e., by E1 or ~ E1) is incomplete in the sense that it fails to account for (by making explicit) all the non negligibly recurring conditions an individual might experience. The global level interpretation then becomes the only available level of interpretation as it tacitly includes all possible variables (including both E1 and ~ E1) that influence organismal success. This decisiveness can, however, blind us to important insights (e.g., negative phenotypic correlations) that appear only upon finer-grained but incomplete partitioning. Opting in this direction threatens to restrict the perceived explanatory and predictive role of fitness in areas of population biology that are concerned with developing a more comprehensive understanding of evolutionary dynamics on the basis of shorter (ecological) timescales.

The deficiency of Static Conditional Fitness should not be taken as demonstrating the impossibility of reconciling the static and dynamic conceptions of fitness. We prefer to view its defects as instructive. Abrams, perhaps more so than anyone else in the philosophical literature on fitness, has clearly articulated how to make the notion of fitness explanatorily (causally) relevant without resort to propensities. Alongside Pence and Ramsey's revamped propensity interpretation, his "trait type" (i.e., organism-environment) conception of fitness is, arguably, the most sophisticated competing causalist account of fitness to date. These two accounts occupy opposing extremes on a continuum of philosophical explications that are committed to the idea that fitness is indispensable to population biology because it is causally efficacious as well as static. Bearing in mind the still pervasive criticism of PIF, the inadequacy of Abrams's account helps reveal that the choice between Static Conditional Fitness and Block Fitness comes with unavoidable epistemic costs. What takes shape is an epistemic trade-off between the predictive accuracy and explanatory (causal) relevance of static fitness measures, one that apparently arises whenever one shifts the assessment of population dynamics from short-term to long-term timescales (for more on this subtle point, see Doulcier et al., 2021). In light of this, philosophers of biology should "pick their poison" carefully when it comes to their preferred notion(s) of fitness.

Natural selection is defined in terms of differential fitness. Arbitrariness regarding the nature of fitness has implications for the nature of selection. If the presence, direction, and magnitude of natural selection are subject to definitional or descriptive fiat, it becomes increasingly difficult to assume that selection is the process most responsible for adaptative evolution. The focused criticism we lodge against Abrams's account of fitness has potentially important implications for a cluster of related issues in evolutionary theory. One of these issues is the nature of inheritance, as some have already argued (see Bourrat, 2019; Charbonneau & Bourrat, 2021). If a given character or trait type may be described more or less finely (e.g., "red" vs. "crimson"), then alternative grains of description can lead to different degrees of resemblance between parents and offspring (e.g., "red" ➔ "red" vs. "scarlet" ➔ "crimson"). The more familiar problem of conventionalism in the levels of selection literature is similarly concerned with how best to describe the levels and units of selection (see, e.g., Kitcher et al., 1990; Waters, 2011; Sober, 2011; Bourrat, 2021). The "grain of description" problem we highlight is also linked to the more general reference class problem in probability (see Hájek, 2003, 2007), which some view as an insurmountable epistemic problem. We believe that such pessimism may be premature and urge more philosophers and biologists to think seriously about this conceptual tangle at the very foundations of evolutionary biology.

Appendix

Statistical measures (e.g., relative fitness inequalities) in Table 1 are based on the following data

Individuals by Trait Type	Number of Offspring	Statistics			
Local Partition E1				**Global Partition (Union of E1 and ~ E1)**	
A	11	Arithmetic Mean for A in E1	10		
A	9				
A	10				
A	9	Arithmetic Mean for B in E1	6		
A	11	Variance for A in E1	0.8		
B	7	Variance for B in E1	0.8		
B	5	Geometric Mean for A in E1	9.96		
B	6	Geometric Mean for B in E1	5.93	Arithmetic Mean for A	8
B	5			Arithmetic Mean for B	8
B	7			Variance for A	4.8
Local Partition ~ E1				Variance for B	4.8
A	7	Arithmetic Mean for A in ~E1	6	Geometric Mean for A	7.69
A	5			Geometric Mean for B	7.69
A	6				
A	5	Arithmetic Mean for B in ~E1	10		
A	7	Variance for A in ~E1	0.8		
B	11	Variance for B in ~E1	0.8		
B	9	Geometric Mean for A in ~E1	5.93		
B	10	Geometric Mean for B in ~E1	9.96		
B	9				
B	11				

Funding Peter Takacs's research was supported under Australian Research Council's Discovery Projects funding scheme (project number FL170100160).

Declarations

Conflict of interest The Authors declare that there are no conflicts of interest.

References

Abrams, M. (2007). Fitness and Propensity's annulment? *Biology and Philosophy, 22*, 115–130.

Abrams, M. (2009a). Fitness 'Kinematics': Biological Function, Altruism, and Organism–Environment Development. *Biology and Philosophy, 24*, 487–504.

Abrams, M. (2009b). What determines biological fitness? The problem of the reference environment. *Synthese, 166*, 21–40.

Abrams, M. (2012). Measured, modeled, and causal conceptions of fitness. *Frontiers in Genetics, 3*, 196.

Abrams, M. (2013). Populations and pigeons: Prosaic pluralism about evolutionary causes. *Studies in History and Philosophy of Biological and Biomedical Sciences, 44*, 294–301.

Abrams, M. (2014). Environmental grain, organism fitness, and type fitness. In G. Barker, E. Dejardins, & T. Pearce (Eds.), *Entangled life: Organism and environment in the biological and social sciences*. Springer.

Ariew, A., & Ernst, Z. (2009). What fitness Can't be. *Erkenntnis, 71*(3), 289–301.

Ariew, A., & Lewontin, R. C. (2004). The confusions of fitness. *The British Journal for the Philosophy of Science, 55*(2), 347–363.

Beatty, J.H. and Finsen S.K. (1989). Rethinking the Propensity Interpretation: A Peek inside Pandora's Box, in What the Philosophy of Biology Is: Essays Dedicated to David Hull, ed. Michael Ruse (Dordrecht: Kluwer Publishers), pp. 17–30.

Bouchard, F., & Rosenberg, A. (2004). Fitness, probability and the principles of natural selection. *British Journal for the Philosophy of Science, 55*, 693–712.

Bourrat, P. (2015). Distinguishing natural selection from other evolutionary processes in the evolution of altruism. *Biological Theory, 10*, 311–321.

Bourrat, P. (2017). Explaining drift from a deterministic setting. *Biological Theory, 12*, 27–38.

Bourrat, P 2019. Natural selection and the reference grain problem". Studies in History and Philosophy of Science Part A. https://doi.org/10.1016/j.shpsa.2019.03.003

Bourrat, P 2021. Facts, conventions, and the levels of selection. Elements in the Philosophy of Biology. Cambridge: Cambridge University Press, 2021. https://doi.org/10.1017/9781108885812

Brandon, R. N. (1978). Adaptation and evolutionary theory. *Studies in History and Philosophy of Science Part A, 9*, 181–206.

Brandon, R. N. (1990). *Adaptation and environment*. Princeton University Press.

Brandon, R. N. (2005). The difference between selection and drift: A reply to Millstein. *Biology and Philosophy, 20*, 153–170.

Brandon, R. N., & Beatty, J. H. (1984). The propensity interpretation of 'fitness'—No interpretation is no substitute. *Philosophy of Science, 51*, 342–347.

Charbonneau, M., & Bourrat, P. (2021). Fidelity and the grain problem in cultural evolution. *Synthese.* https://doi.org/10.1007/s11229-021-03047-1

Doulcier, G., Takacs, P., & Bourrat, P. (2021). Taming fitness: Organism-environment interdependencies preclude long-term fitness forecasting. *BioEssays, 43*, 2000157.

Gillespie, J. H. (1977). Natural selection for variances in offspring numbers: A new evolutionary principle. *American Naturalist, 111*, 1010–1014.

Glymour, B. (1999). Population level causation and a unified theory of natural selection. *Biology and Philosophy, 14*, 521–536.

Hájek, A. (2003). Conditional probability is the very guide of life. In Probability is the very guide of life: The philosophical uses of chance, edited by Kyburg Jr, E. Henry, and Mariam Thalos, 183–203. Open Court,

Hájek, A. (2007). The reference class problem is your problem too. *Synthese, 156*, 563–585.

Kitcher, P., Sterelny, K., & Waters, C. K. (1990). The illusory riches of Sober's monism. *The Journal of Philosophy, 87*, 158.

Kokko, H., Griffith, S. C., & Pryke, S. R. (2014). The hawk–dove game in a sexually reproducing species explains a colourful polymorphism of an endangered bird. *Proceedings of the Royal Society B: Biological Sciences, 281*, 20141794.

Levins, R. (1968). *Evolution in changing environments: Some theoretical explorations*. Princeton University Press.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press.

Metz, J. A. J., Nisbet, R. M., & Geritz, S. A. H. (1992). How should we define 'fitness' for general ecological scenarios? *Trends in Ecology and Evolution, 7*, 198–202.

Mills, S. K., & Beatty, J. H. (1979). The propensity interpretation of fitness. *Philosophy of Science, 46*, 263–286.

Millstein, R. (2016). Probability in biology: The case of fitness. In A. Hájek & C. R. Hitchcock (Eds.), *The Oxford handbook of probability and philosophy* (pp. 601–622). Oxford University Press.

Otsuka, J. (2016). A critical review of the Statisticalist debate. *Biology and Philosophy, 31*, 459–482.

Otsuka, J., Turner, T., Allen, C., & Lloyd, E. A. (2011). Why the causal view of fitness survives. *Philosophy of Science, 78*, 209–224.

Pence, C. H., & Ramsey, G. (2013). A new Foundation for the Propensity Interpretation of fitness. *British Journal for the Philosophy of Science, 64*, 851–881.

Pence, C. H., & Ramsey, G. (2015). Is organismic fitness at the basis of evolutionary theory? *Philosophy of Science, 82*, 1081–1091.

Ramsey, G. (2006). Block fitness. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 37*, 484–498.

Ridley, M. (1996). *Evolution (2nd edition)*. Blackwell.

Rosenberg, A. (1982). On the propensity definition of fitness. *Philosophy of Science, 49*, 268–273.

Rosenberg, A. (1983). Fitness. *Journal of Philosophy, 80*, 457–473.

Rosenberg, Alexander and Frédéric Bouchard. 2015. "Fitness," *Stanford Encyclopedia of Philosophy* (Spring 2020 edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/spr2020/entries/fitness/>

Rosenberg, A., & Williams, M. B. (1986). Fitness as primitive and propensity. *Philosophy of Science, 53*, 412–418.

Scriven, M. (1959). Explanation and prediction in evolutionary theory: Satisfactory explanation of the past is possible even when prediction of the future is impossible. *Science, 130*, 477–482.

Sober, E. (1984). *The nature of selection: Evolutionary theory in philosophical focus*. University of Chicago Press.

Sober, E. (2001). The two faces of fitness. In R. S. Singh, C. B. Krimbas, D. B. Paul, & J. H. Beatty (Eds.), *Thinking about evolution: Historical, philosophical, and political perspectives* (pp. 309–321). Cambridge University Press.

Sober, E. (2011). Realism, conventionalism, and causal decomposition in units of selection: Reflections on Samir Okasha's evolution and the levels of selection. *Philosophy and Phenomenological Research, 82*(1), 221–231.

Sober, E. (2013). Trait fitness is not a propensity, but fitness variation is. *Studies in History and Philosophy of Biological and Biomedical Sciences, 44*(3), 336–341.

Stevens, Lori. 2011. "Selection: Frequency-dependent," in *eLS (Wiley Online Library)*. https://doi.org/10.1002/9780470015902.a0001763.pub2

Tuljapurkar, Shripad. 2013. Population dynamics in variable environments (2nd edition). Springer: Berlin Heidelberg.

Wagner, G. P. (2010). The measurement theory of fitness. *Evolution, 64–65*, 1358–1376. https://doi.org/10.1111/j.1558-5646.2009.00909.x

Walsh, D. M. (2007). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science, 74*(3), 281–303.

Walsh, D. M. (2010). Not a sure thing: Fitness, probability, and causation. *Philosophy of Science, 77*, 147–171.

Waters, K. C. (2011). Okasha's unintended argument for toolbox theorizing. *Philosophy and Phenomenological Research, 82*, 232–240.

Walsh, D.M., Ariew, A., and Matthen, M. (2017). Four Pillars of Statisticalism, *Philosophy, Theory, and Practice in Biology*, vol. 9, no. 1, https://doi.org/10.3998/ptb.6959004.0009.001

Williams, M. B. (1970). Deducing the consequences of evolution: A mathematical model. *Journal of Theoretical Biology, 29*, 343–385.