

Is Organismic Fitness at the Basis of Evolutionary Theory?

Charles H. Pence and Grant Ramsey*†

Fitness is a central theoretical concept in evolutionary theory. Despite its importance, much debate has occurred over how to conceptualize and formalize fitness. One point of debate concerns the roles of organismic and trait fitness. In a recent addition to this debate, Elliott Sober argues that trait fitness is the central fitness concept, and that organismic fitness is of little value. In this paper, by contrast, we argue that it is organismic fitness that lies at the bases of both the conceptual role of fitness and its role as a measure of evolutionary dynamics.

1. Introduction. In a recent paper, Elliott Sober argues that the fitness of individual organisms in the sense usually described by propensity theorists is useless to the actual practice of evolutionary biology. Rather, the crucial sense of fitness for the study of evolution is the fitness of traits, and it is “population-level variation in [trait] fitness”—rather than the absolute value of trait fitness—“that is a causal propensity” (Sober 2013, 337). Indeed, Sober argues that only for variations in trait fitness can a tenable propensity interpretation be constructed; there exists no consistent propensity account of trait fitnesses themselves.

Sober’s argument has much to recommend it. First and foremost, his clarity regarding the distinction between individual fitness and the fitness of traits, as well as the relationship between the two, has been sadly lacking in

*To contact the authors, please write to: Charles H. Pence, Department of Philosophy and Religious Studies, Louisiana State University, 102 Coates Hall, Baton Rouge, LA 70803; e-mail: charles@charespence.net. Grant Ramsey, Institute of Philosophy, KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium; e-mail: grant@theramseylab.org.

†Thanks to Elliott Sober and Carlos Mariscal for comments. Thanks also to audiences at the 2014 Philosophy of Science Association Meeting, the 2014 British Society for Philosophy of Science Meeting, and the Duke Philosophy of Biology Reading Group.

recent literature on fitness.¹ But we will argue here that his central thesis—that individual fitness is broadly irrelevant—is mistaken, and that this mistake arises as a result of confusion over the variety of roles that the notion of fitness plays in evolutionary theory. While trait fitness is the salient concept for some of the roles of fitness, for other uses—and uses in which philosophers are particularly interested—it is individual fitness that is the relevant fitness concept. Sober’s conclusion is thus too hasty; individual fitness remains vital to the practice of evolutionary biology and for the interpretation of evolutionary theory.

Many of the most important uses of fitness fall under two categories. First is what we will call a *metrological* role of fitness—that is, fitness’s role as a quantitative measure in evolutionary studies. Biologists can measure the realized fitness of organisms by tallying such things as their lifetime reproductive success, and they can measure trait fitness by recording trait changes over time.

Second is what we will call the *conceptual* role of fitness—that is, fitness as an element of the causal or explanatory structure of evolutionary theory. It is this sense of fitness to which Abrams appeals when he says that “the kind of fitness relevant to natural selection is fitness of *types*, that is, properties of organisms, since it is types that are heritable and selected for” (2009, 751–52), and to which Pence and Ramsey appeal when they argue that “organismic fitness plays important roles in parts of ecology and evolutionary biology, and is the concept of fitness underlying the [propensity interpretation of fitness]” (2013, 871–72). Here we are considering a deeper, interpretive question about natural selection: fitness either plays some sort of causal or explanatory role in the theory of evolution by natural selection, or it does not—and if it does play a role, then the specifics of that role need to be clarified. It is this role of fitness that we refer to as its conceptual usage.

Keeping this distinction in mind, then, our argument proceeds as follows. In section 2, we argue that there exist three common conceptions of trait fitness—and each of these, in turn, is parasitic on individual fitness, making

1. While the debate over biological individuality and the levels of selection is undeniably relevant to work on the concept of fitness (Bouchard and Huneman 2013), the term ‘individual’ should be taken to be equivalent to ‘organism’ in the following. Further, in his paper, Sober refers to organismic fitness as “token” fitness, while trait fitness is referred to as “type” fitness. We avoid these locutions for several reasons. First, one could construct “type-organism” concepts of fitness. Second, while traits are something like “types” in the sense familiar from metaphysics, they are restricted to particular populations and environments (i.e., we are not interested in the fitness of the type “organism with brown fur,” but in the fitness of the trait “brown fur” within a population, in an environment, at a given time). To avoid these (and other) complications, we will refer exclusively to trait and organismic (or individual) fitness in the following.

individual fitness the fundamental notion of fitness in the conceptual role. In section 3, we argue that in the metrological role, the situation is less clear—there are certainly studies in which trait fitness is the more important concept. But it is, we claim, far from true that, as Sober argues, “evolutionary biology has little use for [individual] fitness” (2013, 336). In a wide variety of examples, we argue, it is indeed the fitness of individual organisms that biologists look to measure, even when they make inferences about the fitness of traits from those measurements. Individual fitness is therefore fundamental in the conceptual role and useful in the metrological role and should thus, contra Sober, by no means be rejected outright.

2. The Conceptual Role of Organismic Fitness. In order to understand the conceptual role of organismic fitness in evolutionary theory, we must know what trait fitness is and how it is related to organismic fitness. We will therefore begin by reviewing the uses of trait fitness in its conceptual role in the philosophical literature. We will then show how these concepts are related to one another and to organismic fitness, finally arguing that organismic fitness lies at the conceptual basis of each of the trait fitness concepts and is therefore at the conceptual basis of the theory of evolution by natural selection.

2.1. Three Concepts of Trait Fitness. We will introduce three definitions intended to capture the core conceptual usage of trait fitness. Nothing in this section should, notably, strike philosophers of biology as particularly surprising or controversial, since these three definitions of trait fitness appear throughout philosophical work on fitness and natural selection.² Further, and importantly, as we will note at the end of this section, these three definitions are often interchanged with one another. Despite the fact that these definitions are often treated as terminological variants, we suggest that they are in fact in profound tension, and their being used interchangeably is deeply problematic.

The first concept of trait fitness holds that the fitness of a trait is the average of the fitness values of the individuals that carry the trait:

(TF1) The fitness of a trait t is equal to the average individual (organismic) fitness values of individuals bearing t .

2. Notably, they also appear throughout—and are used on both sides of—the debate between “causal” and “statistical” interpretations of evolutionary theory. We do not intend anything here to privilege or argue for one of these two positions over the other; these definitions could describe either causally potent or causally impotent concepts.

Commitment to TF1 is widespread and quite explicit. To take one example, Sober (2001) notes a tendency for equivocation between individual and trait fitness. He then asserts, however, that the choice of trait or individual fitness is merely semantic, because the two are related by TF1. That is, “the fitness value of a trait is the average of the fitness values of the individuals that have the trait” (26). Many other authors also explicitly adopt this definition, including Mills and Beatty (1979, 276), Walsh, Lewens, and Ariew (2002, 462), Abrams (2009, 752), and Godfrey-Smith (2009, 21).

Second, spurred by the usage of fitness within population genetics, trait fitness is often definitionally linked to trait dynamics:

(TF2) The fitness value of a trait is a quantity that is, given some model of population dynamics, predictive of the future dynamics of that trait in a population.

This finally lets us cash out some of the value of trait fitness. We want trait fitness to enable us to predict that, in a given population, the fitter traits will, all other things being equal, tend to drive out the less fit.

In biological terms, TF2 is nebulous, since “future trait dynamics” is a multivalent concept. There are countless models connecting fitness to future outcomes, and there are countless future outcomes we might want to observe, from the simple fraction of a trait in a population to times to extinction or fixation. For our purposes, we intend TF2 not to pick out any one of these as privileged, but as a highly general definition of trait fitness: whatever we might think that trait fitness is, it must give us some (reasonably accurate) handle on future trait dynamics. Consider, for example, the way in which trait fitness is defined in the population genetics literature. In the simplest models of population genetics—haploid organisms reproducing asexually in discrete time without overlapping generations—the “Darwinian fitness,” w , may directly provide us with the future proportion at some time t of two competing alleles in a population, p_t/q_t , given their initial proportion (Hartl and Clark 1997, 215):

$$p_t/q_t = w^t \cdot p_0/q_0. \quad (1)$$

In this and many other models of population genetics, the Darwinian fitness is effectively definitionally connected to the changes in allele frequencies over generational time. TF2, when used by philosophers, seems to capture their concern for preserving this usage of fitness in population genetics.

If TF2 does not hold, it is often argued, there is no reason to bother with trait fitness in the first place. A good example here is the work of Ariew and Ernst, who argue that we “employ the concept of fitness when we want to explain why a trait spreads through a population when it does,” and that it is

a condition of the adequacy of a fitness concept that it “enable us to compare the degree to which natural selection will favor the spread of one trait over another, alternative trait” (2009, 290).³

The third concept of trait fitness invokes fitness’s colloquial usage as a description of the “advantage” or “benefit” that an individual organism receives in virtue of possessing a trait:

(TF3) Trait fitness is the reproductive advantage to the individual conferred by possessing the trait.

This definition echoes the original usage of ‘fitness’ in evolutionary theory—the fact that organisms bearing some traits are “better fitted” to their environment than those with other traits (Darwin 1859).

2.2. The Relationship between TF1, TF2, and TF3. Before we consider the relationship between organismic fitness and TF1–TF3, we will briefly consider the relationship between these trait fitness concepts. These definitions are often conflated in the literature, and our analysis here shows that such conflations are deeply problematic.

Consider the pictures of trait fitness invoked by TF1 and TF2. If TF1 is the operative definition of trait fitness, then trait fitnesses, taken to be averages of individual fitness values, are just one of the causal influences responsible for determining future trait frequencies. But now turn to the case of TF2. If a model like equation (1) defines trait fitness, then trait fitness includes the effect of (at least) heritability—future trait frequencies are determined only by current trait frequencies and current trait fitnesses. Trait fitness in the sense of TF1 does not include the impact of heritability, but trait fitness as TF2 does. In many populations, therefore, TF1 and TF2 will result in different values for the fitnesses of traits.

The same argument applies to the relationship between TF2 and TF3. If a trait has a significant benefit to individual organisms, yet is not (or not efficiently) transmitted from parents to offspring, then the TF3-fitness of that trait may be high while its TF2-fitness remains low.

Finally, the relationship between TF1 and TF3 is similarly complex. Consider a trait that constitutes a fairly minor benefit to organisms and the TF3-fitness of which is hence relatively small. If this trait were to occur only in organisms possessing an otherwise extremely fit genetic background, then the TF1-fitness of the trait might nonetheless be quite high. As another example, a novel trait could be instantiated in a sterile individual. In such a case, this trait would have a TF2-fitness of zero, as the only in-

3. Explicit mentions of TF2 also appear in Abrams (2009, 752) and Krimbas (2004, 188).

dividual organism bearing it will have no offspring whatsoever and hence has an individual fitness value of zero. And this would be true regardless of the trait's TF3-fitness value. The average fitness of the individuals bearing a trait can be large (or small), that is, without the effect on individuals being positive and large (or negative and deleterious) in all cases.⁴

It is also noteworthy that the ranges of possible values for these different notions of trait fitness differ.⁵ Individual fitness values can only be positive numbers (an individual cannot have negative fitness), so the TF1-fitness of a trait can only be positive. The TF2-fitness or TF3-fitness of a trait, on the other hand, can clearly be negative—if a trait is declining in frequency within a population, or if it is deleterious to the individual who holds it, then its TF2- or TF3-fitness values, respectively, will be less than zero.

2.3. The Relationship between Trait and Individual Fitness. It is clear, owing to both the extensive use of trait fitness in the literature and the wide variety of ways in which it is defined, that Sober is quite right to argue that trait fitness is an important component of the conceptual foundations of evolutionary theory. But, as noted in the introduction, we take issue with his claim that trait fitness is the conceptually fundamental notion of fitness in evolutionary theory. We will now demonstrate that, for each of the three definitions we offered of trait fitness above, organismic fitness is the conceptually fundamental concept. While trait fitness concepts are valuable, individual fitness serves as the conceptual foundation for all our uses of fitness in evolutionary theory.

Consider first TF3. In order to properly apply TF3 to a particular trait, we need to have a grasp on the appropriate notion of “benefit to the individual.”⁶ How are we to understand such a concept? As mentioned above, many possible “benefits” can be conceived. They all have one important characteristic in common, however—all will involve references to the fitness of individual organisms. Precisely the challenge of developing a model of individual fitness is to determine the way in which various putatively

4. Further instances of this sort can be constructed by appealing to the effects of variance on fitness, as described by Gillespie (1974), or by considering cases of pleiotropy—a pleiotropic trait can have only one TF1-fitness (the average of its varying effects on organisms with different genetic backgrounds), but its TF3-fitness might vary radically across those different organisms.

5. Normalizing these values could, of course, solve this, but this approach is not taken in the literature.

6. One could, conceivably, have a “type”-based notion of TF3, where the discussion of “benefit” was of “benefit to the type” (thanks to Elliott Sober for pointing out this possibility). It is not clear to us, though, that this would resolve the issues raised here: it does not seem plausible that one could somehow discover what the benefit to a type of individual is without clarifying the benefit to token individuals.

beneficial influences should be factored into the overall picture offered by fitness. Importantly, though, it is precisely this work that needs to be performed in order to clarify the notion of “benefit to the individual” that is invoked by TF3. To put the point differently, the work of fully specifying TF3 to the extent that it can actually be used to describe any particular trait will require the construction of a measure of benefit to the individual. This, in turn, just is the construction of a model of individual fitness. However TF3’s invocation of benefit might be cashed out, then, it will ultimately depend on some concept of individual fitness.⁷

The conceptual dependence of TF1 on individual fitness is nearly trivial—if trait fitness simply is the average of individual fitness values, then individual fitness is assuredly the conceptually fundamental notion for TF1. On TF1, trait fitnesses can be defined in terms of individual fitnesses, but the converse is impossible. Similarly, information about individual fitness can derive TF2 values, but TF2 values cannot derive individual fitness values.

The most difficult case is TF2. As Sober noted, TF2-fitness is in fact a fairly heterogeneous property, including such effects as heritability and individual fitness. The question at hand is whether, as for TF1 and TF3, individual fitness also lies at the conceptual basis of TF2. We contend that this is indeed the case. Our argument for this conclusion is that when TF2 is analyzed, individual fitness is one of its core components, but not vice versa. To see this, consider that TF2 is a rate of change in a population. If we ask what underlies this rate of change, the answer will involve several components. If there is immigration, then the immigrants can change trait frequencies. Similarly, emigration can change frequencies, especially if there is a difference in the propensity of different types in the population to emigrate. Mutations and transmission biases, though often small effects, can also change population trait frequencies. All of these factors can change the way in which natural selection operates—but none of them are natural selection, and one of the main causes of trait frequency change (or stability) remains the individual fitness values of the organisms in the population. Although there can be TF2 values in the absence of individual fitness differences, such TF2 values would not indicate an adaptive response. Instead, they are merely due to migration, mutation, and so forth. It is thus true that when we analyze TF2, organismic fitness is not just an important factor but the central factor for understanding the adaptive import of TF2 values.

Now consider individual fitness. Is TF2 at its basis? The answer is no: individuals have fitness values that help lead to TF2 values, but because

7. For a discussion of some of the problems that models of individual fitness need to overcome, see Sober (2001), Abrams (2009), and Pence and Ramsey (2013).

TF2 takes into account population factors like mutation and migration, and because such factors are extrinsic to the propensities of individuals to survive and reproduce, there is no sense in which TF2 lies at the conceptual foundation of organismic fitness. While it is true that organisms are built out of traits, and it is these traits that crucially determine organismic fitness values, it is not true that trait fitness determines organismic fitness values. TF2 and organismic fitness clearly bear an asymmetric relation to one another, and it is organismic fitness that is conceptually primary.

We should pause here to deal with one objection. A response to this discussion of TF2 might run as follows: Of course TF2 is not a complete account of the fitness of traits—we need to include explicit accounts of other properties, such as heritability, population/trait dynamics, and so forth. Once enough of these factors have been considered, only then can we say that we've arrived at a true account of trait fitness.⁸ Our reply to this objection is that it seems to invoke something like a limiting process, where we begin with the limited information offered to us by TF2 and add to it until we have arrived at a “complete” picture. But in what would this complete picture consist? It seems, we claim, that some notion like the concept of “benefit to the individual” invoked by TF3 must be the “target” of the limit, and this would therefore collapse a TF2-notion of trait fitness into one based on TF3. In this case, all the arguments that we deploy with TF3 would then apply.

It is thus clear that, however we choose to define trait fitness, we are left with a notion of trait fitness that fundamentally depends on the concept of individual fitness. As far as the conceptual role of trait fitness is concerned, then, it is the case that individual fitness always stands conceptually prior to trait fitness.

Of course, as mentioned above, the conceptual role is not the only one in which trait fitness features. When Sober argues that “biologists don’t bother with the fitness of Charlie the Tuna, though they may want to discuss the fitness of tuna dorsal fins” (2013, 337), he presumably means that individual fitnesses are of little to no use in the empirical arena, or for what we called the metrological role of trait fitness. It is to this role that we now turn.

3. The Metrological Role of Organismic Fitness. At first blush, it would seem that Sober’s argument against the usefulness of organismic fitness rests on entirely plausible premises. The fitness of organisms is typically inaccessible. This is because “organisms taste of life but once” (Sober 2013, 337). Sober’s argument seems to say that even though organisms have fitness values, unless the values are zero (through infertility, say), we cannot measure them. We saw in the first section that this measurement-focused

8. Thanks to Elliott Sober for offering this response.

(metrological) role of individual fitness can be distinguished from its conceptual role. Because of this, individual fitness can clearly be the conceptual foundation of evolutionary theory even if it is not readily measurable. In this section, however, we would like to address the metrological question. Is it really true that biologists never care about or measure the fitness of Charlie the Tuna?

One excellent resource for gauging the degree to which individual fitness plays a role in evolutionary studies comes from Endler's (1986) classic monograph on the study of natural selection in the wild. In chapter 3, Endler identifies 10 distinct methods for studying natural selection in the wild. These methods vary from method i, which seeks correlations between environmental factors and traits, to method x, which compares optimization models with actual trait distributions. It is clear that for some of the methods, it is traits that are central, not individual organisms and their fitness values. But for at least some of the methods, the fitness of individual organisms plays a central role. Consider method vii, cohort analysis. In Endler's words, "By gathering detailed data on individuals, data can be obtained on survivorship, fertility, fecundity, mating ability, and so on. Data on parents and offspring can also provide information on genetics (condition *c* for natural selection, inheritance). Data are best gathered from individually marked individuals, though some information can be gained by giving all members of the same cohort the same mark" (1986, 81). This method clearly focuses on individual fitness. But in order for method vii to serve as a counterexample to Sober, we will need some sense of how often this method is used in studies of natural selection in the wild.

Method vii is not one that is easy to perform, especially for some taxa. As Endler notes, it "can be the most laborious method" (1986, 81). Does the fact that it is this laborious, however, mean that it is so useless that, as Sober argues, biologists need not (or cannot) bother with attempting to measure individual fitness values? Fortunately for us, Endler took the trouble to conduct a thorough survey of studies directly demonstrating natural selection in the wild. His table 5.1 lists 139 species along with the methods used in the study of each species, as well as the publications that have described these studies. If Sober is right that individual fitness is worthless, we should find that few or none of the studies listed in the table employ method vii. It turns out, however, that a majority of the species (~57%) listed in the table have had natural selection demonstrated in populations via method vii, that is, 79 species mention vii as a method in their studies. Method viii, which also sometimes focuses on individual fitness (though combined together into "age classes" of individuals), is mentioned for 57 species. If we subtract the 18 species whose study has involved both vii and viii, we have a total of 118 species that have been subject to methods vii or viii, 85% of the total. Thus, if we assume that Endler's list is represen-

tative of the kind of studies conducted today, we cannot avoid the conclusion that individual fitness dominates the metrological role of fitness.

On the face of it, then, it seems that biologists wishing to demonstrate natural selection in the wild do care about the fitness of individuals. Charlie the Tuna's fitness is worth measuring, after all. In the previous section we showed that individual fitness is at the conceptual foundation of evolutionary theory, and in this section we have shown that individual fitness plays important metrological roles in many (or even most) evolutionary studies. The claim that individual fitness is useless, then, is difficult to maintain. Is the case closed on individual fitness being the metrological and conceptual foundation of evolutionary biology? Before we can draw this conclusion, we should consider a case study, focusing on just what sort of role individual fitness actually plays in evolutionary studies of the type that Endler catalogued.

Consider a typical method *vii* study, that of Booth (1995). Booth tattooed damselfish in a reef ecosystem and then tracked their fates. By following the outcomes of individual life histories, the study was centered on individual fitness. The determination of individual fitness was not, however, the aim of the study. Rather, Booth was trying to determine the impact of grouping behavior on individual fitness. Is it a fitness advantage to be prone to join groups? And are larger or smaller groups the best ones to join? In terms of our TF1–TF3 framework, we can understand the study as proceeding this way: Individuals are identified, and their fitness values are recorded along with traits of interest (in this case the characters of the groups they belong to). The data from similar individuals can then be averaged, resulting in the TF1-fitness of the traits measured. This average was then used to parameterize models that offered predictions about future evolutionary dynamics (TF2) and also to estimate the impact that various group sizes have on the individual (TF3).

Thus, just as biologists will be more interested in how dorsal fins affect tuna fitness than the fitness of an individual tuna, they will also be more interested in the fitness effect of particular traits (like tending to join large groups) than in the fitness of particular damselfish. In such cases, individual fitness is frequently used as a means of exploring questions about the evolution of traits. But even if this is true, it still does not mean that individual fitness does not play an important role. In fact, we hope to have shown that the fitness of individuals serves as the basis for the demonstration of natural selection in a large percentage of these kinds of empirical studies. This is perhaps to be expected if, as we argued in section 2, individual fitness lies at the conceptual basis of evolutionary theory.

4. Conclusion. We have argued against Sober's contention that individual fitness is useless to the practice of evolutionary biology. While we agree that

trait fitness is sometimes the biologist's sole focus, two facts make Sober's claim incorrect. First, conceptually, each of the three common definitions of trait fitness in fact conceptually relies on the fitness of individual organisms. Organismic fitness thus lies at the conceptual basis of trait fitness. And second, even when biologists are attempting to measure the fitness of traits, they often do so in ways that rely, either tacitly or explicitly, on organismic fitness, making it fundamental as well for the metrological role of trait fitness. Organismic fitness, therefore, is crucial to both the theory and practice of evolutionary biology.

REFERENCES

- Abrams, Marshall. 2009. "The Unity of Fitness." *Philosophy of Science* 76 (Proceedings): 750–61.
- Ariew, André, and Zachary Ernst. 2009. "What Fitness Can't Be." *Erkenntnis* 71 (3): 289–301.
- Booth, David J. 1995. "Juvenile Groups in a Coral-Reef Damselfish: Density-Dependent Effects on Individual Fitness and Population Demography." *Ecology* 76 (1): 91–106.
- Bouchard, Frédéric, and Philippe Huneman. 2013. *From Groups to Individuals: Evolution and Emerging Individuality*. Cambridge, MA: MIT Press.
- Darwin, Charles. 1859. *On the Origin of Species*. 1st ed. London: John Murray.
- Endler, John A. 1986. *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.
- Gillespie, John H. 1974. "Natural Selection for Within-Generation Variance in Offspring Number." *Genetics* 76:601–6.
- Godfrey-Smith, Peter. 2009. *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.
- Hartl, Daniel L., and Andrew G. Clark. 1997. *Principles of Population Genetics*. 3rd ed. Sunderland, MA: Sinauer.
- Krimbas, Costas B. 2004. "On Fitness." *Biology and Philosophy* 19 (2): 185–203.
- Mills, Susan K., and John H. Beatty. 1979. "The Propensity Interpretation of Fitness." *Philosophy of Science* 46 (2): 263–86.
- Pence, Charles H., and Grant Ramsey. 2013. "A New Foundation for the Propensity Interpretation of Fitness." *British Journal for the Philosophy of Science* 64 (4): 851–81.
- Sober, Elliott. 2001. "The Two Faces of Fitness." In *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, ed. Rama S. Singh, 309–21. Cambridge, MA: MIT Press.
- . 2013. "Trait Fitness Is Not a Propensity, but Fitness Variation Is." *Studies in History and Philosophy of Biological and Biomedical Sciences* 44:336–41.
- Walsh, Denis M., Tim Lewens, and André Ariew. 2002. "The Trials of Life: Natural Selection and Random Drift." *Philosophy of Science* 69 (3): 429–46.