



# A New Set of Criteria for Units of Selection

Pierrick Bourrat<sup>1,2</sup>

Received: 29 March 2022 / Accepted: 5 October 2022 / Published online: 24 November 2022  
© The Author(s) 2022

## Abstract

This article proposes two conditions to assess whether an entity at a level of description is a unit of selection *qua* interactor. These two conditions make it possible to (1) distinguish biologically relevant entities from arbitrary ones and (2) distinguish units that can *potentially* enter a selection process from those that have already done so. I show that the classical approaches used in the literature on units and levels of selection do not fare well with respect to either or both of these desiderata.

**Keywords** Additivity · Interaction · Levels of selection · Units of selection · Variance

## Introduction

In his 1966 book, *Adaptation and Natural Selection*, George Williams made the now-famous distinction between a “herd of fleet deer” and a “fleet herd of deer” (Williams 1966, pp. 16–17). He made this distinction to highlight the difference between a benefit at the collective level (i.e., the herd)—that is, a summation of the particles’ benefits that compose the collective (i.e., each deer in a herd)—from a case where this collective benefit is not merely the summation of the benefits that can be attributed to the particles.<sup>1</sup> The fitness of a herd of fleet deer results merely from the fitness of each deer, taken independently. Thus, a herd of fleet deer is also a fleet herd of deer. Yet, when herds have different fitnesses (measured by the number of descendant deer produced after a specific time), we intuitively want to say that selection occurs at the level of the deer rather than the level of the herd. How can we justify this answer in a more systematic way than by merely appealing to our intuitions?

One distinction permitting us to move away (although not wholly) from intuition is to invoke the idea of “cross-level by-product” proposed by Okasha (2006, p. 5). Selection at the herd level, in the case of herds of fleet deer, is only a cross-level by-product of selection occurring at the

deer level. For herd fleetness to be involved in a genuine case of selection at the herd level, selection would have to act directly on the herds as cohesive wholes, to use Hull’s (1980) definition of an interactor.

One primary aim of this article is to provide a rigorous treatment of the notion of a cohesive whole upon which natural selection can occur or, in other words, whether a level of description refers to a level at which there are units of selection—and, if yes, in what sense it does. The notion of a unit of selection has been used in different ways (for a review, see Lloyd 2017). In this article, a “unit of selection” refers to entities that are the “target of selection” or “interactors,” following Hull (1980). I show that many of the classical frameworks for addressing the units of selection problem do not fare well in capturing this account of units of selection on two counts. First, they typically refer to entities that are *actually* rather than *potentially* part of a selection process. I propose, following others, that an alternative characterization of a unit of selection is an entity with the *capacity* to enter a selection process rather than being part of one. In other words, a unit of selection is better seen as a *potential* unit of selection.

Second, I show that previous accounts of the units of selection problem all fall prey to a significant problem, which is appreciated insufficiently in this context—I call this the “arbitrary unit problem.” Put simply, the problem is that these approaches rely on collectives having been identified as units before deciding whether they are units of selection. However, they typically do not provide any rationale

<sup>1</sup> I follow Okasha (2006) in his use of “particle” and “collective” as representing the lower level and the higher level, respectively, in a two-level setting.

✉ Pierrick Bourrat  
p.bourrat@gmail.com

<sup>1</sup> Department of Philosophy, Macquarie University,  
North Ryde, Australia

<sup>2</sup> Department of Philosophy and Charles Perkins Centre,  
University of Sydney, Sydney, Australia

for choosing these entities as units rather than others. Given this, a population of particles could be partitioned into any sort of arbitrary collective entity. As such, partitioning the population of particles into different sorts of collectives could produce a vast number of answers to the question of whether collectives are units of selection in a given setting with no principled way of choosing one of them. In some contexts, this causes no issues; however, in other contexts, such as the context of evolutionary transitions in individuality or the debate over whether some multispecies entities are individuals, the problem is much more significant. By “partitioning,” I mean applying an algorithm to the population of lower-level entities that delineates higher-level entities in a *systematic* way.

In such contexts, it would be valuable to possess a set of tools that permit placing some theoretical constraints on the sorts of collectives that can be considered units of selection or individuals and those that cannot, at a particular level of organization. In this article, I aim to provide those tools.

The article will run as follows. After briefly reviewing several of the significant attempts to characterize a unit of selection and making a few key distinctions, I show that none of these attempts succeeds in characterizing a unit of selection *qua* interactor. The reason is that none is able to both refer to a potential unit of selection and solve the arbitrary unit problem. Then, I propose two conditions—namely, “functional nonadditivity” and “compositional stability”—to distinguish a population of particles where the collectives drawn by an observer are units of selection from a population in which they are not. I show how these conditions can be operationalized using a toy example and discuss their potential use in the context of evolutionary transitions in individuality and the debate over whether some multispecies entities (e.g., biofilms) are individuals.

Before proceeding, I note that for a full evolutionary sequence to occur, at least in paradigmatic cases of evolution, interaction with the environment is an incomplete description of an evolutionary sequence. Evolution occurs as the result of the interplay of two phases in a population: interaction and multiplication (which often involves the transmission of properties from ancestors to descendants).<sup>2</sup> I use the term “multiplication” following Maynard Smith (1983, 1987) and Griesemer (2000). I will restrict my discussion to only the first phase of this sequence and only ask whether an entity interacts with its environment as a

cohesive whole—that is, whether the entity is an interactor, following Hull’s (1980) sense. This is not to suggest that there is nothing to say about the transmission phase of an evolutionary sequence. The question of the nature of multiplication at different levels is certainly interesting; however, it would complexify the analysis provided here. Treatment of reproduction, transmission, and heritability in the context of the units of selection problem and evolutionary transitions in individuality can be found in Bourrat (2019b, 2021a, b).

## Previous Attempts to Characterize Units of Selection

I am not the first with the aim of demarcating genuine units of selection from mere by-products of selection at a lower level. Since the publication of Williams’s book, a number of authors have attempted to make this distinction more systematic. It is only fair to note that this motivation also arose from Lewontin’s (1970) article, “The Units of Selection.” In it, Lewontin argued that for evolution by natural selection to occur, a population only requires three properties: (1) variation between the entities forming the population, which (2) leads to differences in fitness that (3) are heritable. Yet, as noted by Wimsatt (1981, p. 144), one problem with the three conditions proposed is that they do not permit us to distinguish evolution by natural selection occurring at one level due to some cross-level by-products at another level from genuine cases of evolution by natural selection at that level.

Another popular approach to levels of selection is the multilevel form of the Price equation. This equation, proposed by Price himself (see Price 1972a), is derived from its single-level form (Price 1970; Okasha 2006, Chap. 1), the latter of which expresses the total evolutionary change of a character between two times—typically, but not necessarily, generations—in a population as the sum of two terms. The first, referred to as the “selection term,” is the covariance between an entity’s character and its fitness. Assuming there is a causal directional relationship between the character value and fitness, it represents the degree of evolutionary change due to natural selection in the population between the two times. Using Williams’s example, if there is selection for deer fleetness (particle) or herd fleetness (collective), the covariance between deer/herd fleetness and deer/herd fitness will be positive. The second term measures the extent to which, on average, the character of offspring entities deviates from that of their parents multiplied by the fitness of the parental individuals. This term is often called the “transmission bias term.” If the entities of the population reproduce perfectly or there is no deviation from the mean parental character, this term is nil. I do not present the equation here in formal terms since there are a number of introductions

<sup>2</sup> The two phases are conceptual rather than temporal phases, although, in simple cases of discrete generations, they might be temporal. This way of decomposing an evolutionary sequence into two phases—where interaction becomes selection once we are looking at the population level rather than the individual level and eliminate cases of drift—is explicit in the Price approach, among others (see Okasha 2006; Frank and Fox 2020; Bourrat 2019a, for examples).

in the literature (e.g., Frank 1998; Okasha 2006; Bourrat 2021a).

Given that the single-level version of the Price equation can be expressed at any level of organization, one can choose to redescribe a population of particles in terms of collectives. Particles are partitioned into non-overlapping collectives (in any way that will please the observer), and we define the average collective character as the average particle character within each collective, which is also equal to the average particle character in the whole population of particles. From there, we obtain a single-level version of the Price equation at the collective level.

To transform this equation into the multilevel version, one must notice that the transmission bias of the single-level version of the equation at the collective level has the same form as the single-level version at the particle level if we consider that each collective is a population. Thus, if the transmission-bias term of the collective-level version of the equation is replaced by the particle-level version of the equation, we obtain two terms. One term corresponds to the selection occurring between collectives. The other term corresponds to the mean selection occurring within collectives, assuming here that the particles reproduce perfectly so that there is no transmission bias at the particle level. For a formal derivation, see Frank (1998, Chap. 2). For more on the multilevel Price equation, see Price (1972a), Hamilton (1975), and Okasha (2006, Chap. 2).

As numerous authors have argued (e.g., Nunney 1985; Heisler and Damuth 1987; Goodnight et al. 1992; Okasha 2006), the multilevel version of the Price equation, like Lewontin's three conditions, falls prey to the cross-level by-product problem.<sup>3</sup> To see why it does, suppose, following an example used in Sober (1984, p. 260), a population of individuals with different heights and that tallness is favored by natural selection so that, at a later generation, there are more tall individuals and more groups with a taller average height than at an earlier generation.<sup>4</sup> Crucially, assume that there is no effect of the group on individual fitness. In other words, an individual's height does not depend on whether it is in the collective context, which is one way to articulate the notion of the group character "mean height" being a by-product of the height of the individuals composing a group. If one were to apply the

multilevel Price equation to a population of groups with different character values, one could find that there is both a component of collective-level selection and one of particle-level selection. However, recall that we supposed *ex hypothesi* that selection only operates at the individual level. The problem with using this form of the Price equation is that, as exemplified above, the criteria for defining a collective can be chosen on a basis that has no biological relevance and for which we stipulate that there is no effect of the collective on fitness.

Various approaches have been devised to address this problem. The most famous of these is known as "contextual analysis" (Heisler and Damuth 1987; Goodnight et al. 1992; Okasha 2006; Jeler 2014; Earnshaw 2015; McLoone 2015; Bourrat 2016). In the simplest case of contextual analysis, the fitness of an individual in our example is assumed to depend on two characters—namely, its own character (height) and a contextual character (the average height of the group where it is found for a statistical-aggregate collective character). From this assumption, one can write a multilinear regression model with individual fitness as the dependent variable, and individual height and the group average height as independent variables, with their respective strengths measured by partial regression coefficients. Then, this model can be plugged into the single-level version of the Price equation at the particle level. Once this is done, and following several rearrangements and simplifications, we find that the average change in character between two times (i.e., individual height) depends on two components—namely, the variance in individual height in the whole population and the variance in group average height. These two components are modulated, respectively, by the strength of the relationship between each character and fitness (i.e., the two regression coefficients). For a formal derivation, see Okasha (2006, Chap. 3). The term with the variance in individual height is classically interpreted as the "particle-level selection" term, and the term with the variance in average group height as the "collective-level selection" term.<sup>5</sup> In our case, to assume as previously that individual fitness depends solely on individual height (cross-level by-product) is equivalent to a nil partial regression coefficient of fitness on group height, so that the collective-level selection term is nil. Thus, contextual analysis seems to provide a valid method for detecting whether selection occurs at the particle level or the collective level—in other words, to solve the cross-level by-product problem.

In the same spirit as contextual analysis, Wimsatt (1981) and Lloyd (1988) have proposed settling the question of the units of selection problem by relying on a criterion of

<sup>3</sup> That the Price equation, like Lewontin's three conditions, falls prey to the cross-level by-product problem should be unsurprising. As shown by Okasha (2006, Chap. 1), the Price equation can be linked almost straightforwardly to Lewontin's three conditions. One consequence of this is that many of the problems posed for one approach will also be problems for the other.

<sup>4</sup> One could equally use fleetness, but it is easier to measure the average height of a group than the fleetness of deer in herd—hence, I use height and groups for ease of exposition.

<sup>5</sup> For details, see Okasha (2006, Chap. 3).

additive variance in fitness, sometimes called the “additivity criterion.” The criterion, applied to our two-level population of particles that can be partitioned into collectives, relying predominantly on Lloyd’s version, can be summarized as follows:

The collective level is a unit of selection if and only if:

- (1) There is a component of variance in collective fitness that is additive in the population of collectives.
- (2) This additive component of fitness variance does not itself entirely depend additively on variance in particle fitness in the population of particles.

The first clause ensures that selection can occur at the collective level. Without additive variance in fitness, as Fisher’s fundamental theorem dictates, there cannot be selection (Fisher 1930; Price 1972b; Okasha 2008). The second clause ensures that this component of additive variance in collective fitness is not entirely the result of the additive contribution of particle fitness (a cross-level by-product). Another way to make this second point is that additive collective variance in fitness should have a component that depends nonadditively on particle contributions.<sup>6</sup> Take our example of a population of groups of individuals where the group fitnesses depend solely on the fitness of the individuals that compose them. Since all variance in group fitness is additive, the additivity criterion detects—correctly—that the group is not a unit of selection. Again, the cross-level by-product problem seems to be solved.

Brandon (1982; 1990; Brandon et al. 1994) approaches the cross-level by-product problem by exploiting the notion of “screening off” initially proposed by Reichenbach (1956) to characterize causal relationships in probabilistic terms and later developed by Salmon (1971).<sup>7</sup> Suppose a causal chain from  $C_1$  to  $E$  through  $C_2$ , where  $C_2$  depends directly only on  $C_1$ , and  $E$  only on  $C_2$ . If the probability of  $E$  is conditional on  $C_1$ , and  $C_2$  is the same as the probability of  $E$  conditional on  $C_2$  but different from the probability of  $E$  conditional on  $C_1$ ,  $C_2$  screens off  $C_1$ . Formally,  $C_2$  screens off  $C_1$  if and only if:

$$P(E|C_1, C_2) = P(E|C_2) \neq P(E|C_1),$$

where  $P()$  is a probability.

<sup>6</sup> As correctly pointed out by Ketcham (2018, p. 79), Okasha (2006, pp. 116–117) does not represent Lloyd’s additivity criterion accurately.

<sup>7</sup> Brandon’s approach to the problem is actually more complex as it also involves differential reproduction of the unit; however, as I stipulated in the introduction, I do not address the transmission phase of an evolutionary process in this article.

Applied to the units of selection problem, one can propose the following condition for a unit to be a unit of selection<sup>8</sup>:

The collective level is a unit of selection if the fitness<sup>9</sup> of a collective ( $\Omega$ ) conditioning on the properties of its constituent particles ( $p_1, p_2, \dots, p_n$ ) and its collective character ( $Z$ ) is the same as when conditioning on only its collective character but different from when conditioning only on the character of constituent particles. In other words, the collective level is a unit of selection if the collective character screens off particle properties. Formally,  $Z$  screens off  $p_1, p_2, \dots, p_n$  if and only if:

$$E(\Omega|p_1, p_2, \dots, p_n, Z) = E(\Omega|Z) \neq E(\Omega|p_1, p_2, \dots, p_n),$$

where  $E()$  represents an expected value.

More concretely, in a cross-level by-product case, the fitness of a group conditioning on the height of each individual it is composed of and other properties, in addition to its overall mean height, is the same as when conditioning either on only the overall mean height of the group or on the properties of each individual it is composed of. One implication of this is that the character “group height” does not screen off the character “individual height.” For it to screen off this character, the fitness of the group when conditioning on height and other properties of its constituent individuals and its character (average height) should be the same as when conditioning on its character *but different from* when conditioning on the height and other properties of each of its constituent individuals.

Having presented several approaches to the units of selection problem, in the section “[Revisiting the Criteria for Defining Units of Selection](#),” I show that none of these approaches permits us to adjudicate the units of selection problem qua interactor adequately. In the section “[A New Set of Criteria for Potential Units of Selection](#),” I propose two new criteria that permit a remedy for this problem. Before this, in the next section, I define the distinction between actual and potential units of selection, which is essential for assessing the merits of each approach.

## Actual/Potential Units of Selection

In this section, I draw the distinction between a “potential” and an “actual” unit of selection (see Griesemer 2000, p. 70, for a similar distinction). This distinction should be fairly

<sup>8</sup> I follow here the treatment of Sober and Wilson (1994) and Okasha (2006, pp. 121–123) in which conditional expected values are used in lieu of probabilities.

<sup>9</sup> Assuming it can be accurately represented as an expected value.

obvious. A potential unit of selection is an entity type that, once contextualized in a population, *can* be favored or disfavored by natural selection, but is not necessarily so. In other words, it could enter into a selection process but need not do so, due to a lack of variation in the population. An actual unit of selection is an entity type that, once contextualized in a population, *is* favored or disfavored by natural selection. The distinction is, to some extent, similar to the one proposed by Waters (2007) in the philosophy of causation literature and, more particularly, within the interventionist account of causation (Woodward 2003, 2010, 2016; Griffiths et al. 2015), between *actual* and *potential* difference-makers. It can also be traced back to Sober (1984, p. 272), who criticized the analysis of variance (ANOVA)—which is straightforwardly related to the Price equation, contextual analysis, and the additivity criterion—as an adequate method to detect levels of selection in various situations involving a lack of variation. In Sober’s own words: “[i]t is the ANOVA’s obsession with the actual that gets in the way here.”

The distinction between potential and actual units of selection is important in the context of the units of selection problem because, like Sober, I argue that the notion of “unit of selection” ought to refer to a *potential* unit of selection if the status of the objects it refers to are independent of the composition of the population they are part of. For similar claims in the context of evolutionary individuality, see Clarke (2013, 2014, 2016a) and van Gestel and Tarnita (2017).

To see why, suppose a population of entities whose status as units is unquestionable and that they further satisfy Lewontin’s three conditions at the generation  $F_0$ . We could assume, for instance, that the two phenotypes produce a difference in fitness, which is heritable. Suppose now that, for some reason, one phenotype is eliminated or that the environment changes so that the two phenotypes become selectively neutral. As a result, the second of Lewontin’s conditions is no longer satisfied at generation  $F_1$ . Considering that “unit of selection” refers here to *actual* rather than *potential* unit of selection, one would have to conclude that the entity type of this population *was* a unit of selection at the generation  $F_0$ , but no longer is at generation  $F_1$ . However, I claim that, despite not being an actual unit of selection at generation  $F_1$ , this entity type is nevertheless a unit of selection in some relevant sense—namely, it *could enter* a selection process in cases where variation would be introduced in the population.

This reasoning can be reinforced by considering that a “unit of selection” should also be a unit of drift, migration, and mutation. Thus, before determining whether an entity type is a unit of selection, one should ask whether it is a unit at all.

I have argued that considerations of whether selection *does* occur should not drive an answer to the question of the

units of selection. Instead, whether an entity type is a unit of selection should be understood in relation to whether it can enter into a selection process. For this reason, I previously distinguished the notions of potential and actual units of selection.

Stemming from these distinctions is the conclusion that an adequate criterion or set of criteria for a unit to be a potential unit of selection should focus on the properties of entities that are candidate units rather than properties relative to the population where these entities happen to be found.

## Revisiting the Criteria for Defining Units of Selection

With the distinction between a potential and actual unit of selection in place, one reason why some of the approaches to the units of selection problem presented in the section “[Previous Attempts to Characterize Units of Selection](#)” are inadequate appears quite clear. Before stating this reason, I must respond to one potential criticism of my project. It might be argued that these previous attempts do not succeed in demarcating units of selection as I define them because they were not developed with the same concerns as mine. However, this objection may be only *partly* correct since, when the criticisms against the multilevel Price equation were voiced, this was precisely because this approach is unable to distinguish a genuine collective-level component of selection from one that is merely a by-product. Thus, I believe there is sufficient overlap between the way I define a unit of selection and some of the concerns in this literature. That being said, the limitations and criticisms I provide regarding the previous attempts should only be taken as criticisms of the adequacy of these approaches *for a project close to mine*—that is, finding the right level(s) of interaction.

Let us begin with Brandon’s screening-off criterion, which, in some respects, fares the best among the approaches presented above. The first thing to note is that this criterion focuses on some properties of the candidate unit of selection—namely, its absolute fitness—rather than some population-relative properties, such as relative fitness or the additive component of variance in fitness. Thus, using this criterion, whether there is *actual* variation in fitness in the population considered is irrelevant for deciding whether the entity can be a unit of selection. In consequence, the criterion has one of the desirable properties outlined in the previous section—namely, it refers to potential rather than actual units of selection.

This is to be contrasted with the additivity criterion and contextual analysis. Following these two approaches, in any situation where there is no variance in the dependent variable (i.e., fitness), one could conclude two things. First,

one could conclude that collectives are not units of selection—this is despite the fact that, in some cases, from the screening-off criterion perspective, they could be. Or, one could argue that the matter cannot be decided given that the statistical properties of the population to apply contextual analysis or the additivity criterion are not met. This might be the solution favored by Lloyd and the proponents of these and similar methods given that methodological precautions must be taken when using those statistical tools (see Lloyd 1988, p. 75, for a discussion). However, this answer is unpalatable for two reasons. First, a criterion for deciding whether an entity is a unit of selection ought not to depend on our capacity to obtain the right statistical properties. This would be confusing the criterion with our capacity to operationalize it. Second, variance is a property of a population, not the entities that compose it.<sup>10</sup>

Although the screening-off criterion fares better than contextual analysis and the additivity criterion regarding separating potential from actual units of selection, it faces the major problem that it can never be satisfied. To apply Brandon's criterion in the two-level scenario, the collective phenotype is necessarily equal to the sum of the properties of all the collective's constituent particles. Without this property being fulfilled, the mereological supervenience relationship between the two levels is violated. Yet, as noted by Sober and Wilson (1994) in different terms, mereological supervenience renders the criterion impossible to satisfy by definition. This is because, by stipulation, mereological supervenience implies that the collective phenotype is nothing more than the sum of particle properties. Particle properties comprise both *intrinsic* and *extrinsic* particle properties (I discuss these two notions in the next paragraph) that compose the collective. Describing the collective phenotype in terms of particle properties is simply a redescription of the collective phenotype in terms of particle properties.<sup>11</sup>

To satisfy the inequality presented in the section “[Previous Attempts to Characterize Units of Selection](#)” in a two-level scenario, the particles' properties should only refer to a *subset* of all the particles' properties. Perhaps Brandon had in mind the *intrinsic* properties of particles when proposing his criterion? Understood loosely, intrinsic properties are properties of objects that do not depend on the presence and arrangement of other objects (Godfrey-Smith 2009, p. 53). They are *context-independent* properties, as opposed

to extrinsic properties, such as being part of a collective or being at a particular location.<sup>12</sup>

Thus, if my interpretation of Brandon's criterion is correct, his criterion could be reformulated as follows:

The collective level is a unit of selection if the fitness of a collective ( $\Omega$ ), conditioning on all particle-intrinsic properties ( $i_1, i_2, \dots, i_n$ ) and its collective character ( $Z$ ), is the same as when conditioning on only its collective character but is different from when one knows only the intrinsic character of the particles that compose it. In other words, the collective level is a unit of selection if the collective character screens off particle properties. Formally,  $Z$  screens off  $i_1, i_2, \dots, i_n$ , if and only if:

$$E(\Omega|i_1, i_2, \dots, i_n, Z) = E(\Omega|Z) \neq E(\Omega|i_1, i_2, \dots, i_n).$$

Although this revised criterion is an improvement compared to the original one, it suffers from a limitation associated with a second problem. Contextual analysis and the additivity criterion also face this problem. Put simply, the limitation is as follows: the screening-off (revised) criterion, contextual analysis, and the additivity criterion all assume that an observer (which can refer here to a community of researchers) has partitioned a population of particles into collectives. As such, they do not permit discriminating, at a given level of organization, whether the class of units chosen by the observer is a class targeting genuine units—that is, possessing some biological relevance—as opposed to a more arbitrary one. Without making this step principled, there is a risk that this choice is made following the intuitions of the observer, which might be wrong, or that it is defined in a conventional rather than factual way, with potentially different observers or communities using different conventions and talking past each other. Consequently, any tool aiming to address this issue in a principled way would be beneficial—even if, in some contexts, collectives are obvious. A proposal that only solves the cross-level by-product problem, as does contextual analysis, is insufficient.

Yet, the screening-off criterion, once applied, can tell us whether a collective entity is *not* a potential unit of selection—that is, whether a collective-level character is a by-product of a particle-level character. However, it cannot tell us whether the collective entity *is* a potential unit of selection. It does not permit discriminating cases where the collectives drawn by an observer are arbitrary entities from instances where they are not. In fact, if the expected fitness of the collective conditioning on the collective character

<sup>10</sup> For different and related criticisms of the additivity criterion, see Godfrey-Smith (1992), Sarkar (1994), Lewontin (1991), and Sober and Wilson (1994).

<sup>11</sup> Note that this argument is impervious to any empirical considerations. It is merely a methodological assumption that higher-level properties supervene on lower-level ones, following what Sterelny and Griffiths (1999, pp. 138–139) call the “antireductionist consensus.”

<sup>12</sup> For more on the distinction between intrinsic and extrinsic properties and the problems associated with it, see Weatherson and Marshall (2017) and Bourrat (2015, 2017, 2022b, for refinements of Godfrey-Smith's idea when applied to an evolutionary context).

is the same as the fitness of the collective conditioning on solely all the intrinsic properties of the particles composing this collective, the conclusion will be that the collective level *is not* a unit of selection. If, however, the two probabilities are different, there is no guarantee that the collective entities are genuine as opposed to arbitrary entities. This is because the criterion will be satisfied whenever the interactions between particles within a collective drawn by an observer affect the collective fitness. Such situations include cases where the particles are gerrymandered into collectives without any biological relevance (see the section “A New Set of Criteria for Potential Units of Selection” for examples) or in cases where there are no biological reasons to organize particles into collectives, as in many cases of viscous populations or frequency-dependent selection. The problems of gerrymandered units and arbitrary units in viscous populations are a subset of a more general problem I term the “arbitrary unit problem.” The cross-level by-product problem is also a subset of this problem; however, it is distinct from the two others.<sup>13</sup>

Similarly, although I do not show it here (for details, see Bourrat 2021a), contextual analysis and the additivity criterion—although they both permit partially solving the cross-level by-product problem for actual units of selection—nevertheless both fall prey to the two other subproblems of the arbitrary unit problem. Indeed, both assume that higher-level units have already been chosen by the observer.<sup>14</sup>

To summarize so far, the screening-off criterion refers to potential units of selection, while contextual analysis and the additivity criterion refer to actual units of selection. Nevertheless, all three approaches take for granted that collectives have already been defined. Consequently, none of the three methods can assess whether an entity is an arbitrary or genuine unit. Similar conclusions have been reached independently and from a different perspective by Glymour (2017).<sup>15</sup> Glymour suggests that there is no solution to this problem. However, in the next section, I propose the beginning of a solution.

<sup>13</sup> The gerrymandering problem is also encountered in the literature that attempts to define a biological population (see Millstein 2009, 2010; Barker and Velasco 2013; Stegenga 2016; Gannett 2003; Matthewson 2015).

<sup>14</sup> Another problem of contextual analysis specifically (understood following Okasha 2006) is that, even in situations where there are interactions, it will detect a component of collective-level selection that does not exist because the fitness of a particle is assumed to depend on its own character and the character of its collective (which includes the focal particle). This has led Okasha, following Nunnery (1985), to develop an alternative partitioning he calls the “neighbor approach.” However, I will not discuss this issue here.

<sup>15</sup> I thank an anonymous reviewer for pointing this out.

## A New Set of Criteria for Potential Units of Selection

The previous section demonstrated that none of the standard criteria proposed in the literature permits the units of selection problem to be addressed satisfactorily. This is for two primary reasons. First, and most critically, none of the techniques permits deciding decisively whether the collectives chosen are in genuine units upon which natural selection can occur—this embodies the arbitrary unit problem. To understand why this is a significant problem, I follow Millstein (2009) in her view that a solution to the gerrymandered unit problem—which is analogous to one subproblem of the arbitrary unit problem mentioned earlier—in the context of delineating the boundaries of a biological population, is essential since without it what we call selection (and drift) becomes purely arbitrary. She proposes a thought experiment where drawing the boundaries of the population in different ways changes the answer to whether drift or natural selection occurs in the population. She considers this “an unacceptable conclusion for anyone who thinks that selection can explain, as Darwin sought to explain, ‘the mutual relations of all organic beings to each other and to their physical conditions of life’ (Darwin [1859] 1964: 80)” (Millstein 2009, p. 268).<sup>16</sup> Similarly, in the context of units of selection, I propose that it would be unacceptable that the extent to which collective-level selection occurs changes merely because there is no fact of the matter about whether a collective is a genuine one.

Further, a thorough approach to the question of units of selection qua interactor should provide tools that permit us to explain why particles have been partitioned into collectives the way they have. Clarke (2013) proposes that an evolutionary individual, which elsewhere she equates with a unit of selection (see Clarke 2016b), is characterized abstractly by demarcating mechanisms. Yet none of the techniques discussed so far relies on any of those mechanisms; rather, they take the units as given.

The second reason why the techniques surveyed are unsatisfactory concerns only some of them—namely contextual analysis and the additivity criterion. We saw that these two approaches do not permit us to distinguish a potential from an actual unit of selection (assuming we are dealing with nonarbitrary units). They are only able to assess whether selection *does occur* in a given population—rather than whether selection *could* occur. However, I have argued that the latter is the relevant question regarding units of selection

<sup>16</sup> For another related problem, see the example presented in Takacs and Bourrat (2021) where selection can disappear when what is considered the boundary of a population changes.

rather than the former because claiming the contrary would have very unpalatable implications, as detailed above.

In response to the first problem, some might be tempted to argue that whether an entity type is a genuine unit is, in many cases, determined by some relevant biological facts that are easy for an observer or community to detect. For instance, they might argue that boundedness seems relevant here and that whether an entity is bounded is easily observable. I respond that boundedness is a vague concept requiring a precise measure. Further, the relevant type of boundedness for our purpose is not *physical* boundedness, but *causal* boundedness, of which physical boundedness is only one instance.<sup>17</sup> Thus, it is to be expected that, at least in some cases, merely observing a population of entities will not permit partitioning this population into entities that represent potential units of selection. Our tools should provide us with a better justification for partitioning a population of particles in one way rather than another than merely stating that one can observe that it is the correct way or, worse, leaving this decision solely to the intuition of the observer.

Before proceeding further, I must provide some clarification regarding the term “additivity.” The notion of additivity is mathematical. It can be invoked in many contexts and refer to different relationships that satisfy the mathematical property of additivity in some sense. Crucially, within the context of the additivity criterion—this also applies to contextual analysis—the additive component of variance in collective-level fitness (or any character) refers to the relationship between collective *composition*, in terms of particles, and collective-level character. Here, an additive relationship implies that adding one unit of particle character to a collective increases (or decreases) linearly the collective-level character. Significantly, note that if all particles of the population are part of a collective, the only context where the relationship is assessed is within the context of a collective. However, one might mean something different by “additivity”—namely, whether the character contribution of a particle to the collective within the context of a collective is related linearly to the character of this particle when the character of the particle is measured independently of any collective. It might very well be the case that an additive contribution to the collective-level character, when assessed within the context of a collective, does not correspond to an additive contribution when the contribution is assessed with reference to a non-collective context (for a worked-through example, see Bourrat 2021a, Chap. 4). Thus far, the only notion of additivity I have used is the notion of additivity I will hereafter term “contextual additivity.” Crucially,

<sup>17</sup> There is a parallel here with the aim in neo-mechanical philosophy to separate a system into “natural” subunits (for a brief review, see Glennan 2017, pp. 36–44).

additivity in this sense is assessed while particles are always in a collective context. In what follows, I will use a second notion of additivity, which I will term “context-independent additivity” since it refers to additivity in the absence of collective context.<sup>18</sup> An analogous distinction has been made to distinguish physiological from statistical gene–gene interaction (also known as epistasis) in quantitative genetics (for details, see Wolf et al. 2000). (Statistical) additive variance at the organism level can be zero or low in a population while each gene interacts nonlinearly within this organism. Thus, despite a lack of statistical epistasis, there is much physiological epistasis in the population. Similarly, a high level of contextual additivity could be associated with a low level of context-independent additivity.

With these remarks in place, to solve the two problems facing classical approaches to the units of selection problem, I propose the following conditions for an entity to be a potential unit of selection<sup>19</sup>:

In a system composed of lower-level entities, all belonging to the same class of objects, an entity made of lower-level entities is a potential unit of selection if:

1. The character of this entity does not depend purely on context-independent additive contributions of some lower-level character (*functional nonadditivity*).
2. The composition of this entity in terms of the lower-level entities determines its collective characters reliably (*compositional stability*).

Applied to a two-level scenario of particles organized into collectives, the functional nonadditivity condition (1) permits ensuring that the character of the higher-level entity is *not* a cross-level by-product running from the lower to the higher level of organization. In fact, if a collective character has the same value as when each particle character is measured independently and aggregated into a collective, invoking a collective does not add anything to the description in terms of particles only.

Note here that, in their criticism of the additivity criterion, Sober and Wilson (1994) argue that the notion of additivity is irrelevant to the units of selection problem (see also Okasha 2006, pp. 117–119). On this point, I agree only because the additivity they use refers to contextual additivity

<sup>18</sup> Context-independent additivity is a notion close to what Wimsatt (2000) calls “aggregativity.” In Bourrat (2021a, b), I use “aggregativity” instead of “context-independent additivity.” I also use “non-aggregativity” in lieu of “functional nonadditivity.” For our purpose, these terms should be seen as equivalent.

<sup>19</sup> Further elaboration on these criteria are given in Bourrat (2021a, b).



rather than *context-independent* additivity, as I have defined it (see Bourrat 2021a, Chap. 4, for details). The functional nonadditivity criterion I propose is immune to Sober and Wilson's otherwise valid criticism.

The compositional stability condition (2) permits separating cases where the causal interactions between the particles within a collective (chosen by the observer) do not correspond to genuine boundaries. Once operationalized in a population, this condition tells us that, all else being equal, if collectives chosen by an observer with the same composition have different character values, this is evidence that the interactions between the particles in one collective are different from those occurring in the other collectives (assuming here a deterministic setting). Two things should be noted. First, the composition of a collective refers here to lower-level entity measures made independently from the collective. Second, compositional stability should be satisfied for more than one character and in different conditions. If the condition applied to a single collective-level character (drawn by the observer) happens to be contextually additive but functionally nonadditive, the condition will be met for this character—and yet, the collectives might be arbitrary. By having the conditions satisfied for more than one character and in different conditions, the probability that all the characters for a given partitioning are contextually additive and in all environments becomes vanishingly small. However, it is not impossible, and if such a situation were to arise, the criterion would fail.

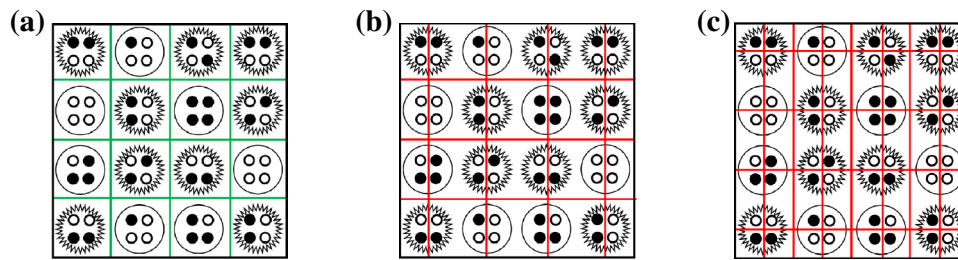
To illustrate how the compositional stability condition can be operationalized, suppose that a boundary (physical or more generally causal) exists between the particles of a collective defined by the observer. Particles on one side of the boundary might interact together but not with other particles on the other side of the boundary, or in a different way. If the observer did not choose—in a biologically meaningful way—to partition the particles into a collective so that the resulting collectives are gerrymandered, one should expect that these causal boundaries are arranged differently in different collectives with the same composition. Consequently, a measure of the collective character should yield different outcomes. On the contrary, if the collectives correspond to genuine units, the causal boundaries (and, consequently, the character) should be the same in different collectives of the same particle composition.

Thus, the second condition permits us to address the arbitrary unit problem when there are interactions between the particles (either because the population is viscous or because there exists one or more scales at which there are collective-level entities). Or, to put it in the terms used by Clarke (2013), it permits us to assess whether the collectives picked out by an observer correspond to units with demarcation mechanisms.

Several remarks should be made at this point. First, the second condition is compatible with the notion of closure of constraints proposed by Montévil and Mossio (2015). In Montévil and Mossio's framework, a biological unit is defined as an entity whose maintenance or persistence is the outcome of a set of causal activities that mutually depend on (or constrain) each other in a cyclic way—hence the label. By selecting an entity where the components realize activities that each depend on one another, one effectively defines an entity that is causally bounded. It follows that if the observer does not partition a population of interacting particles into collectives that exhibit closure of constraints but instead into arbitrary ones, the same type of intervention in terms of particle composition should, on average, lead to different outcomes in terms of collective phenotype. Although I do not show it here, these two conditions can also be related to criteria within the mechanism literature mentioned above, such as “near decomposability” (Wimsatt 1972; Simon 2002), in addition to the recent attempt by Krakauer et al. (2020) to characterize individuality from an information-theoretic approach.

Second, Godfrey-Smith (2008) proposes an analysis of the difference between types of population structure that correspond to genuine cases of collective units from those that do not. He argues that populations organized in genuine collective units are equivalence classes (or close to equivalence classes) when neighbor-structured populations are collective-level entities that are not equivalence classes. A type of entity made of particles is an equivalence class of particles when there exists a binary relation, called an equivalence relation, such as “particle  $x$  belongs to the same higher-level entity as particle  $y$ ,” between the members of the higher-level entity—that is, reflexive, symmetric, and transitive.<sup>20</sup> When the two conditions I propose are satisfied, because they pick higher-level entities that have the same type of causal boundaries, this effectively renders the membership relation between two particles of a higher-level entity as reflexive, symmetric, and transitive. When the population structure is neighbor-structured or viscous, the relationship between two particles of the same neighborhood need not be transitive. Although Godfrey-Smith focuses on the distinction between neighborhoods as opposed to genuine collectives, there is scope to generalize his claim as follows. If a population of higher-level entities is not partitioned into genuine collectives (whether they are a neighborhood or not), the membership relation between any two members of the higher-level entities is not transitive.

<sup>20</sup> The binary relation “ $\sim$ ” for “belongs to the same higher-level entity” is reflexive when “ $x \sim x$ ,” symmetric if “ $y \sim x$ ” then “ $x \sim y$ ,” and transitive if “ $x \sim y$ ” and “ $y \sim z$ ,” then “ $x \sim z$ .”



**Fig. 1** Three different partitionings (different phases and scales) of particles (with two phenotypes, “black” ( $O$ ) and “white” ( $I$ )) into collectives of the same population of particles organized into collectives (with two phenotypes, “smooth” and “spiky”). The spiky phenotype is only expressed if the composition of the phenotype is two black and two white particles. **(a)** Partitioning of particles into collectives at the right scale (four particles) and with the right phase for delineat-

ing potential units of selection at the collective level. **(b)** Partitioning of particles into collectives at the right scale (four particles) but with a wrong phase for delineating potential units of selection at the collective level. **(c)** Partitioning of particles into collectives at the wrong scale (two particles) for delineating potential units of selection at the collective level.

Third, the two conditions I propose do not make mention of fitness. This means that whether the entities satisfying the two conditions are actually undergoing a selection process cannot affect our answer to the question of whether they are units. Thus, these criteria target the units of selection question qua potential units of selection, as required.

Finally, one criticism might be that the two conditions are never satisfied empirically. For instance, to satisfy condition (2) would involve finding two collectives drawn by an observer with the *exact* same character value—which might never be found in nature. To this objection, I respond that the conditions should be regarded as ideal conditions that can be approximated and operationalized empirically. For instance, while two collectives might not have the same value, they might be very similar, which would count (according to some principled reason) as satisfying condition (2), in the same way a statistical test such as a t-test gives us a principled reason to reject the (null) hypothesis that two groups have the same average value for a variable.

With the two conditions now presented, let us examine one way they can be operationalized more precisely with a toy example. Suppose that the population we study is the one presented in the three sub-figures of Fig. 1. Each sub-figure represents the same population of two types of entities (“black” and “white”, with character  $z = 0$  and  $z = 1$ , respectively), but using different ways to partition the particles into collectives. Figure 1a and b are partitionings of the population into entities at the same scale. The only difference between the two is that the phase of the scale has shifted. The notion of phase is used in the context of periodic functions such as sine and cosine. It represents the position at a point in time of these functions. Phase shift is used to characterize a shift on the horizontal axis without any other change in the function properties. Applying this idea to our population Fig. 1a and b show that the grid has been shifted horizontally by one unit of particle entity to the left. This is a two-dimensional representation of a concept that could be

applied to an infinite number of dimensions. Figure 1a and c are partitionings of the population into entities at different scales—namely, four particles and two particles, respectively—while the scales of Fig. 1a and b are the same.

Given these three partitionings at various scales and phases, only the partitioning of Fig. 1a with a scale of four entities and a phase centered on the genuine collective is correct. As I show below, the two conditions proposed above permit us to find these correct scales and phases. Starting with the first criterion, we find that it is satisfied by each of the cases presented in Fig. 1a–c. Indeed, the character of the collective defined by the partitioning in each case is a functional nonadditive function of the particle character. This can be verified by considering that one of the collectives in Fig. 1a has the composition  $0011$  and phenotype  $Z = 1$  (spiky), and another has the composition  $0111$  and phenotype  $Z = 0$  (smooth). Considering that the black and white particles have a phenotype of  $z = 0$  and  $z = 1$ , respectively, when measured independently, the condition is satisfied in this case. (For it not to be satisfied would require the collective phenotype to be proportional to the number of particles of each type when measured independently in each collective, such as the collective  $0001$  having a phenotype  $Z = 0.25$ , because it is composed of one white particle in four particles, while the collective  $0011$  has a phenotype  $Z = 0.5$ , because it is only composed of two white particles in four. In this case, each additional white particle in the collective would increase the collective character by 0.25.) Applying the same reasoning for the two other partitionings, we find that the collectives defined by the partitionings do not have a phenotype that depends linearly (i.e., in a context-independent additive way) on their particle composition, which leads to the same conclusion.

If we now move to the second condition, there are some differences between the three partitionings. In the partitioning of Fig. 1a, if we take collectives with the same composition, they all have the same collective phenotype. For

instance, all collectives with the composition  $0011$  have a phenotype of  $Z = 1$ ; all collectives with compositions  $0111$ ,  $0001$ ,  $0000$ , or  $1111$  have a phenotype of  $Z = 0$ . This is, of course, to be expected since we defined the collective phenotype based on a collective's particle composition, and we applied a partitioning at the scale and phase of the collective we used to define the collective phenotype. If we now move on to the partitioning of Fig. 1b and c and take a collective of four particles in the case of Fig. 1b or two particles in the case of Fig. 1c, we find that the collectives with the same composition vary in their collective character. For instance, in Fig. 1b, there are several collectives<sup>21</sup> with the composition  $0011$ —we find that these collectives can have a phenotype of  $Z = 0$  (smooth),  $Z = 0.5$  (half-smooth, half-spiky), or  $Z = 1$  (spiky). Similarly, in Fig. 1c, we see that the collectives created by the grid with the composition  $01$  can have a phenotype of  $Z = 0$ ,  $Z = 0.25$ , or  $Z = 0.5$ . This reasoning could be generalized to any other partitioning, including partitionings at scales larger than one genuine collective (in our example, larger than four particles).

One could imagine cases with more than two levels and where there would be collectives at different levels of organization, for which the phenotype at that level would be a nonlinear function of the entity immediately at the lower level. Different partitionings at different scales and phases could be tested to delimit each of these collective units at different levels. In cases where the two conditions would not be verified for any scale and any phase, this would be evidence that any population structure existing is not one that individuates particles into collectives. Cases of frequency-dependent selection at a single level of organization fall under this category.

Beyond having at hand a principled approach to define when an entity represents a genuine unit of selection, the approach proposed here could be useful in the context of the origins and nature of individuality. More specifically, the approach could bring new insight to the context of evolutionary transitions in individuality and a closely related question—namely, whether multispecies entities (e.g., biofilms, the gut microbiome, or holobionts) are higher-level individuals.

Abstractly, an evolutionary transition in individuality occurs when individuals at one level of organization start interacting in such a way as to produce larger entities that are then recognized as higher-level individuals (Maynard Smith and Szathmari 1995; Michod 1999; Calcott and Sterelny 2011; Bouchard and Huneman 2013; West et al. 2015; van

Gestel and Tarnita 2017; Black et al. 2020; Bourrat et al. 2022; Bourrat 2022a). A classical example is the transition to multicellularity from unicellular organisms, which would have occurred multiple times in the tree of life, but several others have been proposed (for a discussion, see Bourrat 2011, pp. 6–21). There are typically no issues in recognizing higher-level individuals once a transition is complete. However, such is not necessarily the case when the transition is initiated and mechanisms of demarcations are in formation.

Similarly, it is contentious whether multispecies entities such as biofilms, the gut microbiome, and a holobiont are individuals in their own right. Some argue that they are (e.g., Gilbert et al. 2012; Doolittle 2013; Ereshefsky and Pedroso 2013), while others are more cautious (Clarke 2016b; Skillings 2016; Bourrat and Griffiths 2018).

Applying the criteria of functional nonadditivity and compositional stability could demarcate cases in which a transition is beginning to occur from those in which no transition is occurring. In the context of multispecies entities, the criteria could be deployed to assess whether and the extent to which multispecies entities are individuals.

## Conclusion

In this article, I have argued that none of the approaches for demarcating units of selection in structured populations proposed in the literature is successful. I showed that these approaches face two independent problems. The first problem is faced by some approaches only. These approaches permit us to answer the question of whether a selection process at a given level of organization *does occur* in a population. Still, they cannot inform us whether an entity type *can* enter a selection process. I argued that the second question is the most critical in the context of units of selection. To separate the two questions, I first distinguished the notion of “actual unit of selection” from that of “potential unit of selection.” I also distinguished the question of the levels of selection *once a choice of the units has been made* from the units of selection question per se. The second problem I identified is that in situations where particles interact with one another, none of the approaches classically found in the literature is able to distinguish arbitrary collectives from genuine collectives with real biological significance. To solve these problems, I proposed a set of two conditions that I applied to a toy example to show how they can be operationalized to identify potential units of selection. Further, I highlighted the potential relevance of these criteria in the context of the nature and origins of individuality.

**Funding** This research was supported under the Australian Research Council's Discovery Projects funding scheme (Project Number

<sup>21</sup> Crucially, we are discussing the collective defined by the partitioning, which is represented by a grid on each sub-figure, not the collective defined in our setting, the latter of which represents a “genuine” collective.

DE210100303). The author gratefully acknowledges the financial support of the John Templeton Foundation (#62220). The opinions expressed in this article are those of the authors and not those of the John Templeton Foundation.

## Declarations

**Competing Interests** The author has no relevant competing financial or nonfinancial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Barker MJ, Velasco JD (2013) Deep conventionalism about evolutionary groups. *Philos Sci* 80(5):971–982. <https://doi.org/10.1086/673923>
- Black AJ, Bourrat P, Rainey PB (2020) Ecological scaffolding and the evolution of individuality. *Nat Ecol Evol* 4:426–436. <https://doi.org/10.1038/s41559-019-1086-9>
- Bouchard F, Huneman P (2013) From groups to individuals: evolution and emerging individuality. MIT Press, Cambridge
- Bourke AF (2011) Principles of social evolution. Oxford University Press, Oxford
- Bourrat P (2015) Distinguishing natural selection from other evolutionary processes in the evolution of altruism. *Biol Theory* 10(4):311–321
- Bourrat P (2016) Generalizing contextual analysis. *Acta Biotheor* 64(2):197–217
- Bourrat P (2017) Explaining drift from a deterministic setting. *Biol Theory* 12(1):27–38
- Bourrat P (2019a) Evolution is about populations, but its causes are about individuals. *Biol Theory*. <https://doi.org/10.1007/s13752-019-00329-3>
- Bourrat P (2019b) Evolutionary transitions in heritability and individuality. *Theory Biosci* 138(2):305–323. <https://doi.org/10.1007/s12064-019-00294-2>
- Bourrat P (2021a) Facts, conventions, and the levels of selection. Elements in the philosophy of biology. Cambridge University Press, Cambridge
- Bourrat P (2021b) Transitions in evolution: a formal analysis. *Synthese* 198(4):3699–3731. <https://doi.org/10.1007/s11229-019-02307-5>
- Bourrat P (2022a) Evolutionary transitions in individuality by endogenization of scaffolded properties. *Br J Philos Sci*. <https://doi.org/10.1086/719118>
- Bourrat P (2022b) Unifying heritability in evolutionary theory. *Stud Hist Philos Sci* 91:201–210. <https://doi.org/10.1016/j.shpsa.2021.10.019>
- Bourrat P, Griffiths PE (2018) Multispecies individuals. *Hist Philos Life Sci*. <https://doi.org/10.1007/s40656-018-0194-1>
- Bourrat P, Doucier G, Rose CJ, Rainey PB, Hammerschmidt K (2022) Tradeoff breaking as a model for evolutionary transitions in individuality and limits of the fitness-decoupling metaphor. *eLife* 11(e73):715. <https://doi.org/10.7554/eLife.73715>
- Brandon RN (1982) The levels of selection. *PSA* 1982:315–323. <https://doi.org/10.2307/192676>
- Brandon RN (1990) Adaptation and environment. Princeton University Press, Princeton
- Brandon RN, Antonovics J, Burian R, Carson S, Cooper G, Davies PS, Horvath C, Mishler BD, Richardson RC, Smith K, Thrall P (1994) Sober on Brandon on screening-off and the levels of selection. *Philos Sci* 61(3):475–486. <https://doi.org/10.2307/188057>
- Calcott B, Sterelny K (2011) The major transitions in evolution revisited. MIT Press, Cambridge
- Clarke E (2013) The multiple realizability of biological individuals. *J Philos* 110(8):413–435
- Clarke E (2014) Origins of evolutionary transitions. *J Biosci* 39(2):303–317
- Clarke E (2016a) A levels-of-selection approach to evolutionary individuality. *Biol Philos* 31(6):893–911
- Clarke E (2016b) Levels of selection in biofilms: multispecies biofilms are not evolutionary individuals. *Biol Philos* 31(02):191–212
- Darwin C ([1859] 1964) On the Origin of Species: A Facsimile of the First Edition. Cambridge, MA: Harvard University Press
- Doolittle WF (2013) Microbial neopleomorphism. *Biol Philos* 28(2):351–378. <https://doi.org/10.1007/s10539-012-9358-7>
- Earnshaw E (2015) Group selection and contextual analysis. *Synthese* 192(1):305–316
- Ereshefsky M, Pedroso M (2013) Biological individuality: the case of biofilms. *Biol Philos* 28(2):331–349. <https://doi.org/10.1007/s10539-012-9340-4>
- Fisher RA (1930) The genetical theory of natural selection: a complete variorum edition. Oxford University Press, Oxford
- Frank SA (1998) Foundations of social evolution. Princeton University Press, Princeton
- Frank SA, Fox GA (2020) The inductive theory of natural selection. In: Scheiner SM, Mindell DP (eds) The theory of evolution: principles, concepts, and assumptions. University of Chicago Press, Chicago, London, pp 171–193
- Gannett L (2003) Making populations: bounding genes in space and in time. *Philos Sci* 70(5):989–1001. <https://doi.org/10.1086/377383>
- Gilbert SF, Sapp J, Tauber AI (2012) A symbiotic view of life: we have never been individuals. *Q Rev Biol* 87(4):325–341. <https://doi.org/10.1086/668166>
- Glennan S (2017) The new mechanical philosophy. Oxford University Press, Oxford
- Glymour B (2017) Cross-unit causation and the identity of groups. *Philos Sci* 84(4):717–736. <https://doi.org/10.1086/693873>
- Godfrey-Smith P (1992) Additivity and the units of selection. *PSA* 1992:315–328. <https://doi.org/10.2307/192764>
- Godfrey-Smith P (2008) Varieties of population structure and the levels of selection. *Br J Philos Sci* 59(1):25–50
- Godfrey-Smith P (2009) Darwinian populations and natural selection. Oxford University Press, Oxford/New York
- Goodnight CJ, Schwartz JM, Stevens L (1992) Contextual analysis of models of group selection, soft selection, hard selection and the evolution of altruism. *Am Nat* 140:743–761
- Griesemer JR (2000) The units of evolutionary transition. *Selection* 1(1–3):67–80. <https://doi.org/10.1556/Select.1.2000.1-3.7>
- Griffiths PE, Pocheville A, Calcott B, Stotz K, Kim H, Knight R (2015) Measuring causal specificity. *Philos Sci* 82(4):529–555. <https://doi.org/10.1086/682914>
- Hamilton WD (1975) Innate social aptitudes of man: an approach from evolutionary genetics. In: Fox R (ed) Biosocial anthropology. Malaby Press, London, pp 133–153
- Heisler IL, Damuth J (1987) A method for analyzing selection in hierarchically structured populations. *Am Nat* 130(4):582–602

- Hull DL (1980) Individuality and selection. *Annu Rev Ecol Syst* 11:311–332
- Jeler C (2014) Is there such a thing as “group selection” in the contextual analysis framework? *Hist Philos Life Sci* 36(4):484–502
- Ketcham R (2018) Equivalence, interactors, and Lloyd’s challenge to genic pluralism. In: Jeler C (ed) *Multilevel selection and the theory of evolution: historical and conceptual issues*. Springer International Publishing, Cham, pp 71–98. [https://doi.org/10.1007/978-3-319-78677-3\\_4](https://doi.org/10.1007/978-3-319-78677-3_4)
- Krakauer D, Bertschinger N, Olbrich E, Flack JC, Ay N (2020) The information theory of individuality. *Theory Biosci* 139(2):209–223. <https://doi.org/10.1007/s12064-020-00313-7>
- Lewontin RC (1970) The units of selection. *Annu Rev Ecol Syst* 1(1):1–18
- Lewontin RC (1991) The structure and confirmation of evolution theory. *Biol Philos* 6(4):461–466
- Lloyd EA (1988) *The structure and confirmation of evolutionary theory*. Greenwood Press, New York
- Lloyd EA (2017) Units and levels of selection. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy*. Stanford University, Stanford
- Matthewson J (2015) Defining paradigm Darwinian populations. *Philos Sci* 82(2):178–197. <https://doi.org/10.1086/680665>
- Maynard Smith J (1983) Models of evolution. *Proc R Soc Lond B* 219(1216):315–325. <https://doi.org/10.1098/rspb.1983.0076>
- Maynard Smith J (1987) How to model evolution. In: Dupré J (ed) *The latest on the best: essays on evolution and optimality*, vol 11. MIT Press, Cambridge, pp 119–131
- Maynard Smith J, Szathmari E (1995) *The major transitions in evolution*. Oxford University Press, Oxford/New York
- McLoone B (2015) Some criticism of the contextual approach, and a few proposals. *Biol Theory* 10(2):116–124. <https://doi.org/10.1007/s13752-015-0209-z>
- Michod RE (1999) *Darwinian dynamics*. Princeton University Press, Princeton
- Millstein RL (2009) Populations as individuals. *Biol Theory* 4(3):267–273. <https://doi.org/10.1162/biot.2009.4.3.267>
- Millstein RL (2010) The concepts of population and metapopulation in evolutionary biology and ecology. In: Bell MA, Futuyma DJ, Eanes WF, Levinton JS (eds) *Evolution since Darwin: the first 150 years*. Sinauer, Sunderland
- Montévil M, Mossio M (2015) Biological organisation as closure of constraints. *J Theor Biol* 372:179–191. <https://doi.org/10.1016/j.jtbi.2015.02.029>
- Nunney L (1985) Group selection, altruism, and structured-deme models. *Am Nat* 126:212–230
- Okasha S (2006) *Evolution and the levels of selection*. Oxford University Press, Clarendon Press, Oxford/New York
- Okasha S (2008) Fisher’s fundamental theorem of natural selection: a philosophical analysis. *Br J Philos Sci* 59(3):319–351
- Price GR (1970) Selection and covariance. *Nature* 227:520–21
- Price GR (1972a) Extension of covariance selection mathematics. *Ann Hum Genet* 35:485–490
- Price GR (1972b) Fisher’s ‘fundamental theorem’ made clear. *Ann Hum Genet* 36:129–140
- Reichenbach H (1956) *The direction of time*. University of California Press, Berkeley
- Salmon WC (1971) *Statistical explanation and statistical relevance*. University of Pittsburgh Press, Pittsburgh
- Sarkar S (1994) The selection of alleles and the additivity of variance. *PSA* 1994:3–12
- Simon HA (2002) Near decomposability and the speed of evolution. *Ind Corp Chang* 11(3):587–599. <https://doi.org/10.1093/icc/11.3.587>
- Skills D (2016) Holobionts and the ecology of organisms: multi-species communities or integrated individuals? *Biol Philos* 31(6):875–892. <https://doi.org/10.1007/s10539-016-9544-0>
- Sober E (1984) *The nature of selection*. MIT Press, Cambridge
- Sober E, Wilson DS (1994) A critical review of philosophical work on the units of selection problem. *Philos Sci* 61(4):534–555
- Stegenga J (2016) Population pluralism and natural selection. *Br J Philos Sci* 67(1):1–29. <https://doi.org/10.1093/bjps/axu003>
- Sterelny K, Griffiths PE (1999) *Sex and death: an introduction to philosophy of biology*. University of Chicago Press, Chicago
- Takacs P, Bourrat P (2021) Fitness: static or dynamic? *Eur J Philos Sci* 11(4):112. <https://doi.org/10.1007/s13194-021-00430-0>
- van Gestel J, Tarnita CE (2017) On the origin of biological construction, with a focus on multicellularity. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.1704631114>
- Waters CK (2007) Causes that make a difference. *J Philos* 104(11):551–579
- Weatherson B, Marshall D (2017) Intrinsic vs. extrinsic properties. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy* (fall 2017 edn). <https://plato.stanford.edu/archives/fall2017/entries/intrinsic-extrinsic/>
- West SA, Fisher RM, Gardner A, Kiers ET (2015) Major evolutionary transitions in individuality. *Proc Natl Acad Sci USA* 112(33):10112–10119
- Williams GC (1966) *Adaptation and natural selection: a critique of some current evolutionary thought*. Princeton University Press, Princeton
- Wimsatt WC (1972) Complexity and organization. *PSA* 1972:67–86. <https://doi.org/10.1086/psaprocbienmeetp.1972.3698961>
- Wimsatt WC (1981) The units of selection and the structure of the multi-level genome. *PSA* 1980:122–183. <https://doi.org/10.2307/192589>
- Wimsatt WC (2000) Emergence as non-aggregativity and the biases of reductionisms. *Found Sci* 5(3):269–297. <https://doi.org/10.1023/A:1011342202830>
- Wolf JB, Brodie ED, Wade MJ (2000) *Epistasis and the evolutionary process*. Oxford University Press, New York
- Woodward J (2003) *Making things happen: a theory of causal explanation*. Oxford University Press, New York
- Woodward J (2010) Causation in biology: stability, specificity, and the choice of levels of explanation. *Biol Philos* 25(3):287–318
- Woodward J (2016) Causation and manipulability. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy* (winter 2016 edn). <https://plato.stanford.edu/archives/win2016/entries/causation-mani/>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.