Continuous-time models of group selection, and the dynamical insufficiency of kin selection models

Burton Simon

Department of Mathematical and Statistical Sciences, University of Colorado Denver, 1250 14th St., Denver, CO 80202, United States

HIGHLIGHTS

- A distinction is made between synchronous and asynchronous models of group selection.
- A critique of the Price equation as it is commonly applied to group selection.
- Two-level population dynamics explained in terms of group-level "family trees".
- A new argument for KS≠ GS based on dynamical sufficiency and insufficiency.

ARTICLE INFO

ABSTRACT

Traditionally, the process of group selection has been described mathematically by discrete-time models, and analyzed using tools like the Price equation. This approach makes implicit assumptions about the process that are not valid in general, like the central role of synchronized mass-dispersion and group re-formation events. In many important examples (like hunter–gatherer tribes) there are no mass-dispersion events, and the group-level events that do occur, like fission, fusion, and extinction, occur asynchronously. Examples like these can be fully analyzed by the equations of two-level population dynamics (described here) so their models are dynamically sufficient. However, it will be shown that examples like these cannot be fully analyzed by kin selection (inclusive fitness) methods because kin selection versions of group selection models are not dynamically sufficient. This is a critical mathematical difference between group selection and kin selection models, which implies that the two theories are not mathematically equivalent.

1. Introduction

Steven Pinker's "The False Allure of Group Selection" (Pinker, 2012) along with the responses to it from "The Reality Club", featuring many well known evolutionary theorists, nicely sums up most of the current thinking on group selection. Clearly there is no consensus on which kinds of problems, if any, the concept of group selection can be usefully applied to. Pinker argues that one crucial weakness of the concept of group selection is that the varieties described in words and mathematical models in the literature are not "mechanical" in the profound sense that gene-level natural selection is. (By "mechanical" he means "algorithmic", or "dynamically sufficient" in the usual jargon.) In this article I will argue that group selection is every bit as mechanical as gene-level natural selection, but to see this one must let go of the traditional individual-centered view of the process and embrace a more group-centered view. In general, group selection should be thought of as a dynamical evolutionary "force" that can cause or

speed up the evolution of a genetic trait like "cooperation" by the combined effects of group-level events such as

- *fission*: when a group breaks into pieces that become new autonomous groups,
- *fusion*: when two or more groups merge into one,
- *extinction*: when all the individuals in a group die,
- *group-dispersion*: when some or all of the individuals in one group disperse into the general population, and
- *mass-dispersion*: when the individuals in every group simultaneously disperse, and then recombine into new groups.

This view of group selection in terms of group-level events and long-term outcomes was formally introduced in Simon et al. (2013). It is quite different from the traditional view where individual-level fitness depends (of course) on the composition of the group the individual lives in, but group-level events other than synchronized mass-dispersion do not occur. As we shall see, traditional models do not apply to some of the most important examples of two-level biological processes, so much of the common wisdom on group selection is wrong.

E-mail address: burt.simon@ucdenver.edu

The Reality Club, and the extended research community as a whole, includes a large number of "kin selection partisans" that believe group selection is an unnecessary distraction from real progress in the theory of social evolution, since anything that can be learned from a group selection model can always be better understood using a kin selection (i.e., inclusive fitness) version of the model,[1] e.g., Dawkins (1989, 2012), Frank (2013), Gardner (2008), Gardner et al. (2011), Lehmann et al. (2007), Lion et al. (2011), Marshall (2011), Queller (1992), West et al. (2008), and Wild et al. (2009). However, the Reality Club and extended research community also includes a large number of "group selection partisans" that believe group selection offers a far superior explanation than kin selection for phenomena like the evolution of cooperation in group structured populations, e.g., Boyd and Richerson (1985), Goodnight (1997), Haidt (2010), Henrich and Boyd (2002), Sober and Wilson (1998), Wade et al. (2010), Wilson (2012, 2013), and Wilson and Wilson (2007). Although the details of the arguments between the partisans have changed over the years, this same divide has existed in the research community since the 1960's, which begs the question of how such an important point in evolutionary theory could remain controversial for so long (Okasha, 2001; Lion et al., 2011; Leigh, 2010).

I will argue that much of the disagreement and confusion over the value of group selection is due to a narrow mathematical view of the phenomenon that is widely accepted by partisans on both sides, e.g., Queller (1992), Lehmann et al. (2007), Wild et al. (2009), Lion et al. (2011), Marshall (2011), Gardner (2008), Gardner et al. (2011), Sober and Wilson (1998), Wade et al. (2010), Kerr and Godfrey-Smith (2002), and Okasha (2006). This view of group selection suggests that an analysis based on a detailed "account-ing" of short-term genetic changes in a population is sufficient to understand the process. This "accountant's view" of group selec-tion leads to counter-intuitive and misleading conclusions, such as the well known "mathematical equivalence" between kin selection and group selection, i.e., "$KS=GS$". I will show that in general, kin selection is not equivalent to group selection, except perhaps approximately over short time intervals, such as between group level events. It is ironic and instructive that one of the few things partisans on both sides of the group selection debate agree on turns out to be wrong.

Pinker's criticism of group selection as being non-mechanical applies to many of the models and characterizations of group selection due to group selection partisans, and of course it applies to almost all of the models and characterizations of group selection due to kin selection partisans. However, like all evolu-tionary processes, group selection is clearly a dynamical process. Over time, the effect of group selection is non-zero whenever a population consists of groups of individuals, except in the unrealistic cases where the number of groups and their identities never change (Simon et al., 2013). Although group selection is not always strong, under the right conditions it can be a very potent force.

Evolutionary history is a complicated dynamical process (to put it mildly) where the most interesting and important outcomes often occur in unexpected ways over long periods of time. It is therefore unrealistic to think that much can be learned about an evolutionary process like group selection from an analysis based on short-term genetic changes. Although the time scales are different, evolutionary processes are like global weather processes in many ways. They each have short-term behaviors that are relatively easy to predict and characterize by extrapolating the

current state using current data, like wind speeds and air tem-peratures in weather models, and population sizes and fitness measures in evolutionary models. But there are also complicated trends and counter-trends that play out over longer time frames that cannot be reliably deduced without a proper dynamical model. One should therefore expect the analysis of evolutionary processes to require mathematical techniques similar to those used to study atmospheric dynamics, i.e., differential equations, and at a more basic level, stochastic processes. Just as a thunder storm over a desert environment can be predicted in advance by solving the equations of atmospheric dynamics, the evolution of a trait, like cooperation, in an environment of individuals living in groups, can be predicted in advance from the equations of two-level population dynamics. The equations for two-level population dynamics follow from first principles in the same sense that the equations for atmospheric dynamics follow from first principles. Of course, it is not practical in either case to perfectly specify the real world in all its details, but in principle the dynamics can be predicted. Probably the best way to understand the sequence of events that leads to an outcome like a desert thunder storm is by solving and then animating the dynamical equations that predict the weather, as they do on television. Likewise, the best way to understand how and when group selection occurs is by solving and animating the equations of two-level population dynamics. The process of group selection can be seen clearly in animations of two-level population dynamics, just as a storm can be seen forming in an animation of a weather forecast (Simon and Nielsen, 2012).

The primary purpose of this article is to argue that in general, group selection is an asynchronous, continuous-time process, and should be defined and analyzed as one. Therefore, the traditional approach of analyzing discrete-time models of group selection with the Price equation, and then claiming complete generality, will be criticized. Another primary purpose of this paper is to point out why the Price equation cannot be the basis for an argument that $KS=GS$ (or $KS \neq GS$ for that matter). Misuse of the Price equation is probably the primary reason $KS=GS$ is widely believed. The Price equation has little or nothing to say about group selection. The reasons for this will be explained in Section 3. This article is hardly the first to criticize the kin selection partisans' view that every group selection model should be restated as an (equivalent) kin selection model and analyzed that way (Fletcher and Doebeli, 2009; Fletcher et al., 2006; Nowak, 2006; Nowak et al., 2010a,b; Simon et al., 2012, 2013; Traulsen, 2010; Traulsen and Nowak, 2006; Traulsen et al., 2005, 2008; van Veelen, 2009, 2011; van Veelen et al., 2012, in review). Most prior work arguing $KS \neq GS$ shows that for a specific model of group selection, the corresponding kin selection analysis fails, i.e., when inclusive fitness is defined in the usual way, the answers are wrong. This method for showing $KS \neq GS$ can be countered (so to speak) by redefining inclusive fitness on a case by case basis so it always gives the right answer, e.g., Marshall (2011). Although this counter-tactic usually has no scientific validity (van Veelen, 2011) it would still be useful to find an argument for $KS \neq GS$ that does not rely on specific examples. The present article uses a different tactic to attack $KS=GS$. It will be shown here that continuous-time group selection models are dynamically sufficient (the dynamical equations will be explicitly presented in Section 4), while kin selection *versions* of continuous-time group selection models cannot (even in princi-ple) be dynamically sufficient. No matter how inclusive fitness is defined, no scalar quantity contains enough information to correctly predict the continuous-time dynamics of a two-level population processes. A case can be made that sometimes $KS=GS$ in discrete-time models of group selection, e.g., Kerr and Godfrey-Smith (2002), however, we will see (in Section 5) that

---

[1] The distinction between kin selection theory and inclusive fitness theory, if there is one, is not important in this article, so I will use the terms interchangeably.

in continuous-time, kin selection and group selection models are apples and oranges. The clear distinction between them can be summarized as follows: Kin selection theory is not self-contained (dynamically sufficient) in a multi-level context, while group selection theory is self-contained in a multi-level context, and (of course) only occurs in that context.

2. Synchronous vs. asynchronous models of group selection

Maynard-Smith's (1964) Haystack Model of group selection has been the basic blueprint for many of the subsequent models of group selection in the literature. The basic features of haystack-like models are

- At time $t=0$ a population of "mice" are divided into groups, i.e., one for each "haystack". Individual mice are either Cooperators or Defectors.
- The fitness (expected number of offspring) of each individual in the population can be calculated in terms of its type, and on the composition of its group.
- Each individual contributes its offspring into its group, and then dies.
- The groups simultaneously disperse into a single "cloud" of individuals.
- New groups form out of the cloud of individuals at time $t=1$, and the cycle starts again.

Although models like this may differ in the way fitness is assigned, or in the way the groups reassemble, the salient feature of all haystack-like models is that the only group-level event is mass-dispersion.

Time can be thought of as discrete in haystack-like models since (without loss of generality) all the "action" (individual-level reproduction and deaths, mass-dispersion, and group-re-formation) occurs simultaneously at a discrete set of time points, e.g., $t=1,2,\ldots$. Discrete-time models are appropriate in examples where the entire population's life cycles are synchronized; and some real-world examples are approximately like this, e.g., mice living in haystacks. However, in many other examples, discrete-time models are inappropriate because the life cycles are not synchronized. Of the five group-level events, fission, fusion, extinction, group-dispersion, and mass-dispersion, defined above, only mass-dispersion lends itself to discrete-time models. The other events occur to only one group (or two in the case of fusion) at a time, and since fission, fusion, extinction, and group-dispersion can occur at any time, to any group, and in any order, a continuous-time model is necessary to properly capture the dynamics of the process.

There are many examples of two-level population processes in nature that feature asynchronous group-level events, e.g.,

- Hunter–gatherer tribes are subject to fission and extinction events. Fission events occur when tribes get too big, or perhaps too uncooperative. Group extinction events may occur due to battles or plagues.
- Host/parasite models include group-dispersion and extinction events. Group dispersion occurs when some or all of the parasites are expelled from a host organism, while extinction events occur (for example) if the host dies or fights off the infection.
- Simple multi-cellular organisms are subject to fission, fusion, and extinction events. Fission events occur if an organism breaks into viable pieces. Fusion occurs if two organisms stick together and function as one. Extinction occurs if the organism is eaten or dies.

There are also two-level population process associated with "major transitions", like the emergence of the first replicating cells. For example, the "Stochastic Corrector Model" of early life (Szathmary and Demeter, 1987; Grey et al., 1995), features fission and extinction events.

In fact, *none* of these examples just described feature synchronized group-level events like mass-dispersion, so discrete-time models are not appropriate for any of them. Conclusions about group selection based on discrete-time models cannot be assumed to hold in any of these examples.

2.1. What does dynamically sufficient mean?

It does not appear that "dynamically sufficient" has been rigorously defined in the literature, but the term does seem to be used in a consistent manner. A model is dynamically sufficient if the evolutionary trajectories can (at least in principle) be found unambiguously in terms of the basic model variables, by either an exact mathematical solution, or by an exact (stochastic) simulation or numerical calculation. In other words, a model is dynamically sufficient if it is "self-contained". Not all models are dynamically sufficient. In some haystack-like models the process does not iterate: the process stops as soon as the first cloud of individuals forms, because there is no way to unambiguously determine the statistical properties of the next generation from the model. Those models determine only the "direction of change" in the overall level of cooperation, and are not dynamically sufficient. For haystack-like models, Kerr and Godfrey-Smith (2002) argue that the key to dynamical sufficiency is a set of rules for assembling the next generation of groups from the cloud of dispersed individuals produced by the previous generation, but this cannot be a general definition of the term, since it assumes that the population's life cycles are synchronized. If a discrete-time model is dynamically sufficient then it may be possible to determine the long-term outcome of the process, e.g., Matessi and Jayakar (1976) and Kerr and Godfrey-Smith (2002). Even without dynamical sufficiency it is often tempting to try to determine an equilibrium configuration by (essentially) setting the state at $t=0$ equal to the state at $t=1$, i.e., a static analysis (Frank, 2013). However, without dynamical sufficiency this approach is sometimes difficult to justify mathematically. No such problems exist for continuous-time models of group selection. The state-dependent group-level and individual-level event rates all have to be specified, but the result is always dynamically sufficient. The long-term (equilibrium) outcome is found by solving the equations. The dynamical equations are discussed in Section 4.

There is a crucial distinction between two-level population models where the primary group-level event is synchronized mass-dispersion, and two-level population models primarily driven by asynchronous group-level events like fission, fusion, extinction, and group-dispersion. If there are mass-dispersion events at times $t=1,2,\ldots$, and no other group-level events, then no information about the population dynamics in the continuous interval $t \in (k, k+1)$ is needed except the overall change in the fractions of Cooperators and Defectors. In these cases, the accountant's view of the process is sufficient. Under suitable conditions, these haystack-like models can be fully analyzed using variables corresponding to either individual-level (inclusive) fitness measures, or group-level fitness measures (e.g., average individual-level fitness in the group), implying that $KS=GS$ in these cases. (But even in haystack-like models $KS=GS$ is not guaranteed, van Veelen et al., 2012). The important point is that in general, two-level population processes are not haystack-like. They feature asynchronous group-level events, and therefore must be analyzed with continuous-time models. As we will see, individual-level fitness measures like inclusive fitness, no matter how they are

defined, do not contain enough information to specify self-contained dynamical equations for continuous-time models. Thus $KS \neq GS$ whenever asynchronous group-level events are important in the process.

2.2. MLS1 and MLS2

Examples of two-level selection are often categorized as either MLS1 or MLS2 (Damuth and Heisler, 1988; Okasha, 2006). Roughly speaking, in MLS1 the fitness of a group is the average fitness of the individuals that live in the group, or equivalently, the growth rate of the group. On the other hand, in MLS2 the fitness of a group is the number of offspring groups it produces. Some models are clearly either MLS1 or MLS2; for example, haystack-like models are MLS1. In the next section it will be useful to keep in mind that the two-level version of the Price equation is designed for MLS1 models, while MLS2 models use the one-level version of the Price equation (Okasha, 2006).

However, the two-level population models described here lead to models of group selection that are *neither* MLS1 nor MLS2, so those concepts will not be important in the main discussions here. By taking the long-term view of an evolutionary process, as we do here, the whole idea of "fitness" becomes circular. In general there is no way to determine long-term fitness (some measure of long-term evolutionary success) of a biological entity without solving the dynamical population model first. Here, the basic model data includes no measures of fitness. Instead, the data is composed of state-dependent rates that various biological events (births, deaths, fission, fusion, etc.) occur at. Although measures of fitness can be derived from these rates (by solving the resulting equations), they are not fitness measures themselves. In general when group-level events are asynchronous, there is no unambiguous way to define MLS1 or MLS2 (or any other) fitness measures. Over which time-frame(s) should fitness be measured? Every answer yields a different meaning and value of fitness.

3. The Price equation cannot be applied to asynchronous models of group selection

Price (1970, 1972) discovered an interesting way to write the average genetic change in a population of individuals over a generation (or some other appropriate time unit). The formula decomposes the total change into the change due to "selection" and the change due to other factors.[2] The Price equation has a mathematical structure that allows it to be iterated hierarchically, so an analogous formula can be derived for the average genetic change in a population of individuals that live in groups, e.g., from Gardner (2008),

$$\overline{w}\Delta\overline{g} = \mathrm{Cov}_i(w_i, g_i) + E_i(w_i \Delta g_i). \tag{1}$$

The Price equation for groups of individuals decomposes the average genetic change over the whole population, $\Delta\overline{g}$, into the change due to group selection (the "covariance" term) and the change due to individual-level selection (the "expected value" term). When the covariance term is positive and the expected value term is negative, the interpretation is a "conflict" between the two levels. When $\Delta\overline{g} > 0$, group selection partisans say the increase is due to group selection, while kin selection partisans offer an explanation in terms of inclusive fitness. In either case, this phenomenon is summed up nicely by the well

known quote from Wilson and Wilson (2007), "Selfishness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary." As we shall see, the crucial issue is *how* altruistic groups beat selfish groups. The Price equation (1) only sees when altruistic groups grow faster than selfish groups, i.e., MLS1, which can (but does not always) help cooperators in the long run. The Price equation misses the most important and most interesting ways that cooperation can win in the long run.

The Price equation is a mathematical identity, and one of its celebrated strengths is its wide range of applicability. One hears that there is no imaginable evolutionary scenario where the Price equation does not apply and say something interesting and important. Needless to say, this is an overstatement, and in fact, the Price equation (1) is already in over its head with two-level population processes that feature asynchronous group-level events like fission, fusion, extinction, and group-dispersion. For the purpose of studying group selection, the Price equation has a number of limitations:

1. The Price equation is not predictive. It is not a model. It is an after-the-fact way to represent population-change data that it does not produce itself. It relies on a separate mathematical model, or a lab experiment, to supply the population-change data. The Price equation is simply an algebraic identity, so there is no actual (new) information content in it. Furthermore, the way the Price equation is usually written (1) is misleading since the "covariance" and "expected value" terms are not really statistical quantities.
2. In general, the length of the time step for which the Price equation is valid is limited, since the basic population structure (e.g., the number and identities of the groups) is not allowed to change during the time step.
3. The Price equation implicitly assumes that "group-level fitness" is the average individual-level fitness of the group members, i.e., MLS1. This is simply the growth rate of the group, which is not an accurate measure of fitness when there are asynchronous group-level events.

The first limitation is relevant in every application of the Price equation, not just group selection. It is explained in depth in van Veelen (2005) and van Veelen et al. (2012), so the reader is referred to those articles for a robust critique of the Price equation and its champions. The remaining points address specific issues associated with the mathematical structure of the Price equation (1) that make it inappropriate for studying group selection models.

The Price equation is not predictive (dynamically sufficient) *because* it is not a model. An infinite regress appears when trying to derive the rate of change of the mean fitness from the Price equation, i.e., the change in the mean fitness depends on second moment quantities like the covariance, the changes in the second moment quantities depend on third moment quantities, and so on. This problem of dynamical insufficiency is well known, and is usually considered unimportant by those that favor characterizing group selection with the Price equation (including partisans on both sides), e.g., Frank (2012). The goal of the analysis, they say, is to study general properties of the process itself, and to make comparisons, not to predict long term outcomes. But surely determining the long-term outcome ought to be an important goal of the analysis. By studying the population changes that occur during small time intervals under the microscope of the Price equation, the analyst misses the forest for the trees. In fact no short-term analysis can provide answers to basic questions like "Can cooperation evolve *and thrive* by group selection?" The Price equation can sometimes be rewritten as a condition for the

---

[2] There are other ways to represent genetic change in two-level population structures, such as a technique called "contextual analysis" (Okasha, 2006; Goodnight, 2012) but we focus here on the Price equation since it has been used for decades to study and characterize group selection.

average level of cooperation in the whole population to increase or decrease in the short-term, like Hamilton's rule, Gardner (2008), but it is silent on cooperation's long-term fate. Depending on the initial conditions, it is not unusual for cooperation to increase at first, and then later die off; or conversely, to get off to a bad start but eventually thrive (Simon et al., 2012, 2013). A short-term analysis will get things wrong in both cases. The short-term "direction of change" cannot be reliably extrapolated very far into the future, as every student of calculus knows. A dynamical analysis is necessary to predict long-term outcomes, e.g., extinction of Cooperators and/or Defectors, or a stable equilibrium with Cooperators and Defectors coexisting. Furthermore, a dynamical analysis also provides all the intermediate short-term changes as byproducts, so a long-term dynamical analysis is obviously preferable to a short-term analysis in every sense. As I will explain in the next section, a dynamical analysis is always possible (at least in principle) for any two-level population process satisfying certain minimal structural assumptions. The claim made by some kin selection partisans that a static (equilibrium) analysis of a group selection model is preferable to a dynamical analysis, e.g., Frank (2013), is therefore wrong. Their reason for preferring a static analysis is that by setting their sights lower (no dynamics) they can presumably make their models more general. However, statics is never preferable to dynamics if the dynamical equations are mathematically tractable, as they are in the cases of continuous-time models of group selection.

The second limitation of the Price equation listed above involves the limited kinds of population changes that can be properly represented. Without considerable modification, the Price equation does not allow the basic structure of the population to change during the analyzed time step. Both terms on the right side of (1) involve sums over a collection of groups, $i \in G$, where the set of groups, G, is implicitly assumed to be the same at the beginning and end of the time step being analyzed. The Price equation therefore only applies to examples of two-level population processes where the only things groups can do between time steps is change in size and composition. Group-level events like fission, fusion, extinction, and group-dispersion, change the number and identities of the groups, throwing a monkey wrench into the Price equation. The unit of time for the Price equation therefore has to be small enough so that no group-level events occur during the time step. (It is often suggested that the set G could include "dummy" groups that are empty at the beginning of the time interval, and become nonempty in the event of a new group forming, e.g., from a fission event. But this trick does not work since the calculation of the fitness measures leads to division by zero.) If asynchronous group-level events are common and important in a two-level process (and they typically are) then the Price equation does not apply exactly, even in the short-term. Whether the Price equation can be modified to account for certain asynchronous group-level events and still retain its useful structure is an open question whose answer appears to be no. Even if the Price equation could be modified to account for (say) fission, the solution would be piecemeal since the modification process would need to start over every time a model with a different kind of group-level event (e.g., fusion) came up. In any case, what is clear is that traditional characterizations of group selection based on the Price equation in its usual form (1) do not apply to examples where asynchronous group-level events are important, except perhaps during the time intervals between the events. This is a very serious limitation, and throws into question all the general claims about group selection that are directly or indirectly based on the Price equation.

The third limitation of the Price equation approach to group selection is another consequence of its basic mathematical structure. Eq. (1) implicitly assumes a special form of "group fitness"; namely

w_i, the average fitness of the individuals in group i, or equivalently, the ith group's total growth rate. This sort of fitness measure is appropriate in haystack-like models, but in general "Price fitness", i.e., MLS1, says very little about the future prospects of a trait like cooperation. For example, in Simon et al. (2013), a model of the evolution of simple multi-cellular organisms is studied where the groups (the multi-cellular organisms) with the most Cooperators ("sticky" cells in that case) have the *lowest* Price fitness (i.e., all else equal, sticky organisms do not grow as fast as less-sticky organisms), and in addition, sticky cells are at a disadvantage within each organism (due to a lower birth rate); yet stickiness establishes itself in the end, its success due entirely to the long-term effects of the group-level events, fission, fusion, and extinction. To see that Price fitness can be misleading in another context, consider again an example where multi-cellular organisms are the groups. This time, organisms are made of a variety of different kinds of cells, one of which is cancerous, and causes tumors that eventually kill organisms that contain them. An organism with a fast-growing tumor is obviously not fit; however, w_i is maximized by a fast-growing tumor, since it is the average growth rate of *all* the organism's cells that determines Price fitness. The Price equation does not see into the future when the tumor finally kills the organism, along with itself (a group level event); it only sees the current rapid growth of the organism. One might argue that this example is really MLS2 since the tumor decreases the number of offspring the organism can have. However, it is not clear how to count the number of offspring of an organism that reproduces by asynchronous fission events. (How much time, or how many "generations" do you wait?). In any case, there is no way to determine this sort of fitness without solving the dynamical first, so the one-level Price equation is of no use in this example either.

If there are no group-level events, i.e., if the same set of groups exist forever, then the Price equation is valid. However, since the Price equation does not apply to two-level population processes with asynchronous group-level events, which are common in nature, it follows that opinions on group selection based on its traditional mathematical characterization and analysis are not well grounded. A characterization of group selection, featuring all the important group-level events, and based on long-term evolutionary outcomes, is needed to properly evaluate its status as an evolutionary force.

4. A continuous-time view of group selection

Group selection is a phenomenon that occurs in the context of a two-level evolutionary process, so in order to understand the process of group selection, one must first understand two-level population dynamics. In order to fully appreciate the diagrams and animations of two-level population dynamics in this section, it is first necessary to understand what the "motion" in them represents, which means understanding how the state of a two-level population process is represented, and how it changes. This means keeping track of the effects of individual-level and group level events as they occur (asynchronously) in time. In all that follows it is assumed that there are no mass-dispersion events in the models, since they cause inherent discontinuities in the two-level population dynamics, and therefore do not allow an analysis by differential equations. Models featuring any combination of fission, fusion, extinction, and group-dispersal, (and other similar group-level events) are within the present framework. But, to keep things as simple as possible in the present discussion, a simple generic model of the evolution of cooperation will be used to illustrate the concepts. In the generic model, the only group-level events are fission and extinction, and the only individual-level events are births and deaths. This section is designed to be

self-contained. However, readers interested in more general versions of the model or in the formal derivations and (numerical) solutions of the dynamical equations used here, are referred to Simon et al. (2012) and Simon and Nielsen (2012). The proof of a limit theorem connecting the stochastic and deterministic models is in Puhalskii and Simon (2012), and the formal definition of group selection used here is justified in Simon et al. (2013).

The present model features a population of groups, where each group consists of Cooperators and Defectors. The Cooperators and Defectors play public goods games in their groups. The payoffs to Cooperators and Defectors from their public goods games are used to determine their (instantaneous) birth rates. The (instantaneous) death rates for Cooperators and Defectors are equal to each other within each group, and depend on the size of the group. The individual-level birth and death rates therefore change in time as the states of the groups change. At the group level, the fission and extinction rates depend on the state of the group, and the state of the whole population. See Simon (2010) and Simon et al. (2013) for specific examples of models like these. The state of a group is assumed to depend only on the numbers of Cooperators and Defectors in the group (e.g., the state does not depend on the individuals' ages, the physical location of the group, etc.), so in the present model a group is specified by (x,y), where x is the number of Cooperators and y is the number of Defectors. A group in state (x,y) is called an (x,y)-group. The dynamical state of the environment is represented by

$\theta_t(x,y) = $ number of (x,y)−groups in the environment at time t.

Quantities of interest associated with the model can be calculated in terms of the state of the environment. For example, the number of groups at time t, the number of Cooperators at time t, and the number of Defectors at time t are (respectively)

$$G(t) = \sum_{x,y}\theta_t(x,y), \quad N_c(t) = \sum_{x,y}x\,\theta_t(x,y), \quad \text{and} \quad N_d(t) = \sum_{x,y}y\theta_t(x,y).$$

Fig. 1 shows an example of the state of some generic environment as a function of time, t. Each trajectory in the figure corresponds to a group, so $G(t) = 6, t \geq 0$. The axes in the figure measure the number of Cooperators and Defectors in the groups, so a dot at (x,y) corresponds to an (x,y)-group. Thus, groups near the x-axis are mostly Cooperators, and groups near the y-axis are mostly Defectors. At $t=0$ (the dark dots), $N_c(0) \approx 10+10+30+30+50+50 = 180$ and $N_d(0) \approx 2+20+2+20+2+20 = 66$. The compositions of the groups in the environment change in time due to births and deaths of Cooperators and Defectors, so the dots "move". Fig. 1 is based on the kinds of within-group population dynamics that occurs when the birth and death rates of Cooperators and Defectors depend on how they fare in a public goods game. Other assumptions lead to different trajectories, but the "counterclockwise" trajectories in the figure, due to Defectors winning in each group, are typical of these sorts of models. Here, Cooperators eventually disappear in every group due to the Defectors' higher birth rates, and the all-Defector groups settle at their equilibrium size.

Of course, in a real-world two-level population process, the long-term outcome in Fig. 1 will typically not occur. Groups are not immortal, and so they will not necessarily reach their final equilibrium state. Long before a group reaches internal equilibrium it may break apart (fission) or die (extinction). Without group-level events in the model, cooperation is doomed to extinction, as Fig. 1 shows; but with group-level events there is at least a possibility that somehow cooperation can thrive in the environment. If cooperation does somehow thrive, it is due to group selection, since Cooperators go extinct if group-level events are absent from the model. This definition of group selection is introduced and justified in Simon et al. (2013).
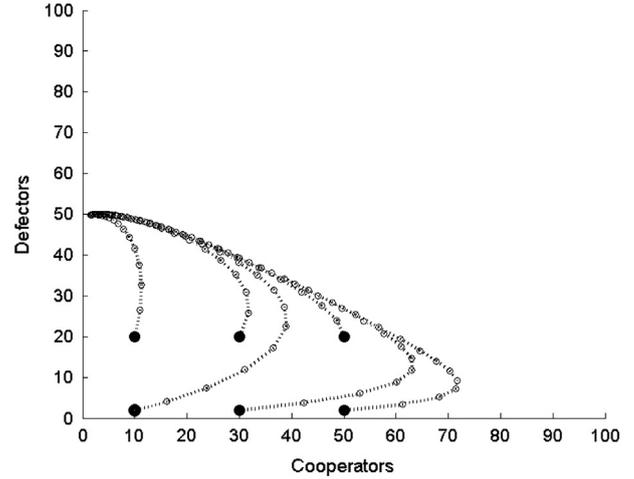

Fig. 1. The state of the environment and how it changes in time. At time $t=0$ there are six groups with states given by the dark dots. The composition of each group changes in time; the spaces between the light dots are identical time intervals. Assuming the groups never fission or die of extinction, the trajectories continue until Cooperators are extinct and the groups consist of 50 Defectors. The counter-clockwise motion reflects the fact that Defectors are favored in every configuration. The smooth trajectories are from the "hybrid solution" of the dynamical equations, explained in Simon and Nielsen (2012).

To understand how group selection can occur through asynchronous group-level events, consider what unfolds over time starting with a single group. Fig. 2 shows a "family tree" of a group and its "progeny" over time, where fission (births) and extinction (deaths) occur. The figure shows three (asynchronous) "generations". Starting at $t=0$ (Fig. 2a) the original group grows and then fissions, yielding two offspring groups "siblings" that then start to grow. In Fig. 2b (which starts where Fig. 2a leaves off) we see that the less cooperative sibling group soon dies of extinction, while the more cooperative sibling continues to grow, and later fissions itself. When the scenario illustrated in Fig. 2b leaves off, one of the surviving "grand-offspring" groups is significantly more cooperative than the original group, and the great-grand-offspring (not shown) could include groups that are more cooperative still. And this occurs *despite* the counter-clockwise trajectories of each member of the family tree. By imagining dozens or hundreds of inter-dependent group-level family trees like this occurring simultaneously, the reader is visualizing the process of two-level population dynamics. Sample animations of this process can be found on YouTube.[3] There is a more extensive catalog of animations of two-level population processes in Simon and Nielsen (2012).

In this article I am concerned with the kind of group selection that causes or assists the evolution of individual-level traits, like cooperation. This sort of group selection can often be thought of as selection for a certain group-level trait. A group-level trait in this context is any well defined function of the state of the group, i,e, $T(x,y)$ is the value of the trait associated with an (x,y)-group. The kinds of group-level traits that cannot be deduced solely from the distribution of individual-types in the group, such as cultural traits, are outside the present discussion. The focus in the present discussion is on a particular group-level trait, $T(x,y) = x/(x+y)$, i.e., the fraction of the individuals in the group that are Cooperators. When a fission event occurs, the fractions of Cooperators in the offspring groups are necessarily correlated in some way with the

---

[3] An animation of the basic stochastic (Markovian) process for a model of the evolution of cooperation can be found at http://www.youtube.com/watch?v=pqT6dJFb6OA&feature=relmfu.
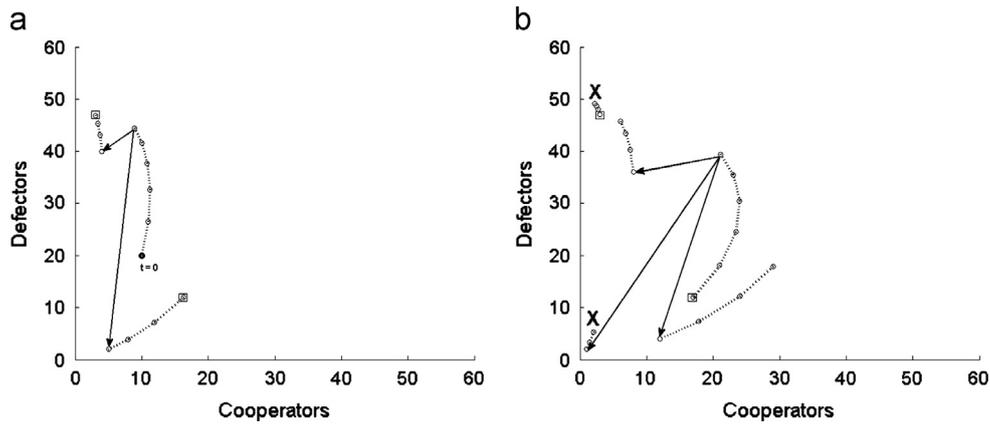
Fig. 2. A group-level "family tree" starting with a (10,20)-group at $t=0$. The curves representing the group life histories are the same sort as in Fig. 1. The arrows reflect a group fission event, and the **x**'s correspond to group extinctions. (b) begins where (a) leaves off (small squares) at about $t=8$. When (b) leaves off at $t=16$, two "grand-offspring" groups are alive, and one of them is significantly more cooperative than the original group, despite the counter-clockwise drift every family member experiences due to Defectors outcompeting Cooperators within them. This counter-clockwise drift is how groups "age".

fraction in the parent group at the time of the fission. For example, if a group with $T(x,y)=0.40$ randomly breaks into pieces, the pieces will typically have an average of about 40% Cooperators as well. Thus group-level traits like T are heritable over fission events. Different fission rules will lead to different kinds of correlations between parent and offspring, however, some variation around the parent's trait value (perhaps a lot of variation) is inevitable. Thus, variability of the trait among offspring, a necessary condition for an evolutionary process, is also present in this context. And finally, if (say) more-cooperative groups are less likely to die of extinction than less-cooperative groups, then the more-cooperative offspring groups from a fission event are more likely to "grow up" and fission later themselves, e.g., Fig. 2. This is group-level selection for more-cooperative groups. The evolution of more-cooperative groups at the group level coincides with (and arguably causes) the evolution of more Cooperators at the individual level.

Of course, within the groups Defectors are favored. There is selection at the within-group (individual) level, which causes the counter-clockwise motion. Due to the within-group population dynamics, the state (and therefore the trait) of a group changes in time, as can be seen in Figs. 1 and 2. The trait value of a group at the time of its birth and at the time of its fission or death will typically not be the same. (But they will obviously be correlated.) Group-level traits are therefore different in kind from individual-level (genetic) traits, which do not normally change during the lifetime of the individual.[4]

Thus, to gauge the strength of group selection in this model the fundamental question is: To what extent does the variability among the pieces resulting from fission events, along with the selective advantage enjoyed by more cooperative pieces, offset the counter-clockwise motion corresponding to Defectors winning in every group? In Fig. 2 we see that (at least in principle) it is possible for more-cooperative groups to emerge despite the counter-clockwise drift. In general it is very difficult to guess how a two-level population process will unfold over time, so the dynamical equations have to be solved.

4.1. Two-level population dynamics from first principles

In this sub-section we move from models where populations are integer-valued to models where populations are real-valued. One way to think about this is to assume the true populations are

very large, and that x and y are "scaled" so they take values in a continuum. The nature of this scaling, and the derivation of the general equation for continuous-time, two-level population dynamics is formally discussed in Simon (2010) and Puhalskii and Simon (2012). Another way to think about the real-valued (continuous) model is that it is simply an approximation of a integer-valued process, i.e., a population of x is rounded to the nearest integer. A close examination of the way the dynamical equation was derived in Simon (2010) reveals that its basic form does not depend on the details of the example being studied. A general continuous-time model of a two-level population process is specified by

1. A collection of "events" that can occur, e.g., births, deaths, migrations at the individual level, and fission, fusion, extinction at the group level.
2. Rules (typically statistical) that specify how the state of the population changes when an event occurs.
3. Rate functions that specify how often (on the average) the various events occur, given the current state.

These rates that events occur can be interpreted as stochastic rates (yielding a discrete-state, continuous-time Markovian model) or as deterministic rates (yielding a continuous-state model based on differential equations). As was just mentioned, the differential equations in the continuous-state model can be thought of as large-population limits of the Markovian model, and in certain cases these kinds of limits have been justified rigorously (Champagnat et al., 2006; Puhalskii and Simon, 2012; Luo, 2013). The derivation in Simon (2010) shows that in the continuous-state case, the dynamical state variable, $\theta_t(x,y)$, obeys an equation of the form

$$\frac{\partial \theta_t}{\partial t} + \frac{\partial(\theta_t \alpha_c)}{\partial x} + \frac{\partial(\theta_t \alpha_d)}{\partial y} = g_t, \tag{2}$$

no matter what the details of the example might be. The functions $\alpha_c(x,y)$ and $\alpha_d(x,y)$ are essentially (instantaneous) fitness functions for Cooperators and Defectors, which characterize the individual-level population dynamics within the groups, and the function, $g_t(x,y)$, characterizes the sum of the group-level effects. Based on the rigorous proofs of similar (but perhaps less complicated) large-population limits in the papers just mentioned, it is reasonable to conjecture that the PDE (2) can be similarly justified by taking the proper limit of the Markovian model of two-level population dynamics. Eq. (2) governs two-level population dynamics in the same sense that the equations of atmospheric dynamics governs the weather. If the initial population and the event rate functions are specified (approximately) correctly,

---

[4] The evolution of a group-level trait, can therefore be thought of as a kind of Lamarckian process.

and the initial populations are large enough, then (2) will correctly predict the ensuing population dynamics. Even when populations are relatively small, the continuous approximation imitates the discrete model very well, as can be seen in animations of the dynamics. Numerical methods for solving and animating the PDE can be found in Simon and Nielsen (2012). Sample animations of two-level population dynamics can be found on YouTube.[5] There is a catalog of animations of two-level population dynamics in Simon and Nielsen (2012).

5. Kin selection and group selection are apples and oranges

The widespread belief that group selection and kin selection are just different ways of looking at the same phenomenon appears to be accepted equally by kin selection partisans and group selection partisans. It is probably the root cause of much of the longstanding and unnecessary confusion in the field. $KS=GS$ has attained the status of a fundamental mathematical result in evolutionary theory; but a search of the literature reveals no general theorem. In fact, there are already demonstrations that $KS \neq GS$ in the literature (Hauert and Imhof, 2011; Simon et al., 2012; Traulsen, 2010; van Veelen et al., 2012, in review) that make the point by constructing examples where a kin selection analysis of a group selection model, using standard definitions of inclusive fitness, gives the wrong answer. Kin selection partisans sometimes adjust the definition of inclusive fitness (or relatedness) to get around these problems when they come up. Although this practice usually has no scientific validity (van Veelen, 2011) it is undeniable that in every example there will exist one or more values of inclusive fitness that would make the numbers "add up" to the right answer. (Assuming the kin selection model satisfies some simple continuity conditions, the existence follows from the Intermediate Value Theorem.) It is difficult to find a counter-example to $KS=GS$ if inclusive fitness is simply defined to be whatever number makes things work! The approach here, which refines an argument outlined in Simon et al. (2013), is different and is not subject to that defensive tactic. Basically, the argument is that asynchronous group-level events make it impossible (even in principle) to compute the time-dependent values of inclusive fitness that would produce the correct evolutionary trajectory, *unless* the two-level dynamical equation (2) has been solved first. But this means that kin selection versions of a continuous-time group selection models are not dynamically sufficient (since they are not self-contained), and therefore cannot be mathematically equivalent to group selection models, which *are* dynamically sufficient.

Before proceeding, let us be clear about what is meant by "KS and GS are apples and oranges". Saying that $KS \neq GS$ does not imply that there are no examples where the theories are interchangeable. Kerr and Godfrey-Smith (2002) demonstrate equivalence for certain haystack-like models, and van Veelen (2009) and van Veelen et al. (in review) are very careful to point out that in certain special cases the results are correct when using an inclusive fitness approach to a group selection model, even in continuous time. However, in order for $KS=GS$ to be true, there have to be *no* counter-examples to the claim. As it turns out, counter-examples appear to be the norm in this context, not the exception.

Before one can say that two mathematical formalisms are equivalent, one needs to carefully define them in a way that they can be compared. In most cases it is not at all clear how a model of group selection, phrased in terms of two-level population dynamics, i.e., Eq. (2), can be translated into a model of kin selection, which is phrased in terms of inclusive fitness measures. It seems to be taken for granted that someone sufficiently good at kin selection modeling can always find a way to transform a given group selection model into an equivalent kin selection model. Although there are examples in the literature where a group selection model was arguably rewritten as a kin selection model, e.g. the Lehmann et al. (2007) reformulation of the Traulsen and Nowak (2006) model, nobody has described a recipe that is guaranteed to translate a group selection model into a kin selection model, like (say) the recipe that translates a calculus problem described in rectangular coordinates to an equivalent problem in polar coordinates. Kin selection partisans think there must be such a recipe based on the Price equation. They argue that since the genetic change over a time step in a model of group selection can always be represented with the two-level Price equation (1), and since the terms appearing in that context can be algebraically rearranged into the Price equation used in the one-level (inclusive fitness) context, the equivalence must be completely general (Gardner, 2008; Lehmann et al., 2007; Lion et al., 2011; Marshall, 2011; Queller, 1992; West et al., 2008; Wild et al., 2009). However, this argument does not work since the Price equation cannot be applied to examples where asynchronous group-level events, like fission, fusion, extinction, and group-dispersion, are featured in the model. As was shown in Sections 2 and 3, some of the most interesting and important examples of two-level population processes feature these kinds of group-level events, so the Price equation cannot be the basis for an argument that $KS=GS$. This observation alone should be enough to deflate the alleged generality of $KS=GS$.

However, the fundamental reason that $KS \neq GS$ has nothing to do with the Price equation. It has to do with dynamics. Continuous-time group selection models are dynamically sufficient since they can be analyzed by the PDE (2). On the other hand, kin selection versions of those group selection models, which encode everything (the effects of individual-level and group-level events) into inclusive fitness measures, are not dynamically sufficient. Although this claim may not be obvious, it should not be surprising. If inclusive fitness is to somehow take group-level effects into account, how can it know what those effects will be without first solving the two-level dynamical equations? Perhaps the best way to understand the problem is by analogy.

5.1. The "inclusive forceness" theory of spring-mass dynamics

Imagine an apparatus featuring a spring with one end fixed, and an object (a "mass") attached to the other end. Suppose the object moves vertically along the z-axis, and let $z(t)$ be the position of the mass at time t. By convention, the position of the non-fixed end of the spring when it is at rest without the mass attached is at $z=0$. In the usual formulation of spring-mass dynamics based on Newton's law, $f = ma = mz''$, the position satisfies

$$mz'' + p(z,t)z' + qz = e(z,t), \tag{3}$$

where m is the mass of the object, $p(z,t)z'$ is the magnitude of the frictional (velocity dependent) forces on the object when it moves past z at time t, e.g., forces due to air resistance or mechanical damping; and qz is the magnitude of the force on the object due to the spring when it is stretched (or contracted) to z, i.e., from Hooke's Law. The right hand side term, $e(z,t)$, accounts for external forces on the object at time t, e.g., forces due to gravity or the physical motion of the whole apparatus (Serway and Jewett, 2003). The point I am trying to make here can be made even in the simple case where $p(z,t) = 0$, but in any case we assume m, $p(z,t)$, q, and $e(z,t)$ are known, so (3) can be solved (using standard numerical techniques if necessary), yielding the position of the object as a function of time, indefinitely into the future.

---

[5] An animation of the solution of the continuous dynamical Eq. (2) for a model of the evolution of cooperation can be found at http://www.youtube.com/watch?v=87UAHkCK1qQ&feature=plcp.

Now suppose there are certain scientists (inclusive forceness partisans[6]) who think that the concept of an external force is unnecessary when studying spring-mass systems. These scientists believe that every force external to the spring itself can, and should, be analyzed as a kind of frictional (or perhaps "anti-frictional") force. In this view, spring-mass dynamics are analyzed by equations of the form

$$mz'' + \hat{p}(z,t)z' + qz = 0, \tag{4}$$

where $\hat{p}(z,t)$ is the "inclusive forceness", which takes into account all forces on the object except the one described by Hooke's Law. By rearranging terms, we see that if one defines

$$\hat{p}(z,t) = p(z,t) - e(z,t)/z'(t), \tag{5}$$

then Eq. (4) has exactly the same solution as (3). This "proves" that if inclusive forceness is defined by (5), then (3) and (4) are mathematically equivalent. External force is therefore an unnecessary concept for understanding and predicting the dynamics of spring-mass systems.

Needless to say, there must be something wrong with the inclusive forceness partisans' reasoning, and indeed there are a number of serious problems with it. First of all, there is a problem with (5) when the object changes direction, i.e., when $z'(t) = 0$. At these points in time, the inclusive forceness is undefined, and so attempting to use (4) to determine the position of the object after a directional change is futile. Thus, inclusive forceness theory only works in the time intervals between direction-changing events. However, an even more serious problem is that inclusive forceness models are not dynamically sufficient. The magnitude of the inclusive forceness term, $\hat{p}(z,t)$, cannot be computed without knowing the object's velocity, $z'(t)$, and the only way to find $z'(t)$ is to find $z(t)$ from (3) first. In other words, Eq. (4) is not self-contained. Of course, once (3) has been solved there is no need to solve (4). Obviously (3) is the correct way to analyze spring-mass systems, and inclusive forceness theory is only appropriate in the special cases where there are no true external forces. The two formalisms for studying spring-mass systems are *not* mathematically equivalent, since one (the usual formulation which keeps frictional and external forces separate) is more generally applicable than the other.

5.2. Back to group selection

Just as inclusive forceness partisans are wrong when they claim that external forces are unnecessary to explain and understand spring-mass dynamics, kin selection partisans are wrong when they claim that group selection is unnecessary to explain and understand two-level evolutionary dynamics. Since they believe $KS = GS$, kin selection partisans must believe that the effects of group-level events on individual-level population dynamics can be correctly taken into account by properly defining inclusive fitness. In other words, the group-level effects accounted for by g_t in (2) can be somehow "absorbed" into the individual-level fitness terms, α_c and α_d, yielding an equation involving only individual-level fitness measures, $\hat{\alpha}_c$ and $\hat{\alpha}_d$, i.e.,

$$\frac{\partial \theta_t}{\partial t} + \frac{\partial(\theta_t \hat{\alpha}_c)}{\partial x} + \frac{\partial(\theta_t \hat{\alpha}_d)}{\partial y} = 0. \tag{6}$$

As with the inclusive forceness example, there is a case to be made that this "consolidation" is possible. After all, the inclusive forceness term $\hat{p}(z,t)$ in (4) can, in a sense, be found by solving (5). Similarly, the inclusive fitness terms $\hat{\alpha}_c$ and $\hat{\alpha}_d$ in (6), designed to

make g_t unnecessary, can be found by solving

$$\frac{\partial(\theta_t \hat{\alpha}_c)}{\partial x} + \frac{\partial(\theta_t \hat{\alpha}_d)}{\partial y} = \frac{\partial(\theta_t \alpha_c)}{\partial x} + \frac{\partial(\theta_t \alpha_d)}{\partial y} - g_t. \tag{7}$$

Of course there is no free lunch in either case. In order to calculate the inclusive fitness terms, $\hat{\alpha}_c$ and $\hat{\alpha}_d$ that appear in (6), Eq. (7) has to be solved first. But since (7) involves θ_t, it is necessary to solve (2) before solving (7). Eq. (6) is therefore not self-contained. The analogy is complete. Inclusive forceness theory fails because it is only valid between direction-changing events, and (more importantly) even in those intervals where it is valid, it is not self-contained and therefore not dynamically sufficient. Likewise, inclusive fitness theory is only valid in the time intervals between group-level events, and the time-dependent values of inclusive fitness cannot be determined unless (2) is solved first. Any analysis of a continuous-time two-level population model based solely on inclusive fitness cannot be dynamically sufficient. Group selection is therefore mathematically distinct from kin selection.

6. Summary and final thoughts

In this article I argued that some of the most important examples of two-level population processes, like hunter–gatherer tribes, host/parasite populations, and simple multi-cellular organisms, are not governed by synchronized group-level events like mass-dispersion. Instead, they are governed by asynchronous group-level events like fission, fusion, extinction, and group-dispersion. This means that these real-world examples of group selection are outside the domain of the Price equation. Since the arguments for $KS = GS$ are usually based on applications of the Price equation to discrete-time models, a general mathematical equivalence between kin selection and group selection has never been demonstrated, and in fact the claim is false. Continuous-time models of group selection are very different from haystack-like models. In particular, while it is possible to argue that $KS = GS$ in certain haystack-like models, it is essentially never true that $KS = GS$ in continuous-time models. The reason is that kin selection versions of continuous-time group selection models are not dynamically sufficient, and the reason they are not dynamically sufficient is that inclusive fitness cannot be computed in those models without solving the group selection model first.

The belief that $KS = GS$ may be the root cause of much of the long-standing confusion and partisanship in this field. If nothing else, it has stifled mathematical research on group selection by implying (among other things) that there is no need for good models of two-level population dynamics. The continuous-time models described in this article are clearly models of group selection, and since they imply $KS \neq GS$, the general mathematical equivalence of kin selection and group selection is an untenable proposition. The correct way to analyze continuous-time models of group selection is with (2), which is a self-contained equation that can be solved numerically, and even animated, yielding a complete picture of group selection processes. It is a mistake to fold the effects of asynchronous group-level events in a two-level population process into inclusive fitness measures and then analyze the result as a kin selection model. The resulting model will lack dynamical sufficiency, so nothing can be learned about its long-term behavior. Furthermore, unless (2) is solved first, the inclusive fitness measures will be miscalculated, so even the short-term behavior will be predicted incorrectly. Of course, once (2) is solved, the kin selection model of the process is superfluous.

In conclusion, it may seem paradoxical that belief in $KS = GS$, which seems to imply some sort of agreement, is actually a contributor to the partisanship in the research community, while acceptance of $KS \neq GS$ has the potential to reduce the partisanship.

---

[6] This tongue-in-cheek parody of inclusive fitness theory and its partisans is meant to be instructive, not offensive.

If $KS \neq GS$, then for a given example the best model is (usually) either KS or GS, depending on whether there are asynchronous group-level events or not, which is usually fairly obvious. There is much less to argue about. On the other hand, if $KS = GS$ then the choice of which to use becomes a philosophical problem, and strong opinions can result. Kin selection is well suited for one-level population processes (subject to certain caveats, van Veelen, 2009), as well as certain synchronized two-level population processes, where its domain overlaps to some extent with group selection. But when asynchronous group-level events, like fission, fusion, extinction, and group dispersion, are important players in a two-level population process, kin selection theory is not a viable mathematical approach to studying the dynamics or long-term outcome of the process. Kin selection theory is not dynamically sufficient. However, since equations like (2) can provide a full dynamical analysis of a two-level population process, including the long-term outcome, the group selection approach is the proper method of analysis.

Acknowledgments

References

Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago.

Champagnat, N., Ferriere, R., Meleard, S., 2006. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. Theor. Popul. Biol. 69, 297–321.

Damuth, J., Heisler, I.L., 1988. Alternative formulations of multi-level selection. Biol. Philos. 3, 407–430.

Dawkins, R., 1989. The Extended Phenotype. Oxford University Press, USA.

Dawkins, R., 2012. Group Selection is a Cumbersome, Time-Wasting Distraction. Responses to Pinker. From the Reality Club ⟨http://edge.org/conversation/the-false-allure-of-group-selection⟩.

Fletcher, J.A., Doebeli, M., 2009. A simple and general explanation for the evolution of altruism. Proc. R. Soc. B 276, 13–19.

Fletcher, J.A., Zwick, M., Doebeli, M., Wilson, D.S., 2006. What's wrong with inclusive fitness? Trends Ecol. Evol. 21, 57–60.

Frank, S.A., 2012. Natural selection IV. The Price equation. J. Evol. Biol. 25, 1002–1019.

Frank, S.A., 2013. Natural selection. VII. History and interpretation of kin selection theory. J. Evol. Biol. 26, 1151–1184.

Gardner, A., 2008. The Price equation. Curr. Biol. 18, R198–R202.

Gardner, A., West, S.A., Wild, G., 2011. The genetical theory of kin selection. J. Evol. Biol. 24, 1020–1043.

Goodnight, C.J., 1997. Experimental studies of group selection: what do they tell us about group selection in nature? Am. Nat. 150, S59–S79.

Goodnight, C., 2012. On multilevel selection and kin selection: contextual analysis meets direct fitness. Evolution 67 (6), 1539–1548.

Grey, D., Hutson, V., Szathmary, E., 1995. A re-examination of the stochastic corrector method. Proc. R. Soc. Lond. B 262 (1363), 29–35.

Haidt, J., 2010. The Righteous Mind: Why Good People are Divided by Politics and Religion. Random House, New York.

Hauert, C., Imhof, L.A., 2011. Evolutionary games in deme structured, finite populations. J. Theor. Biol. 299, 106–112.

Henrich, J., Boyd, R., 2002. On modeling cultural evolution: why replicators are not necessary for cultural evolution. J. Cognit. Cult. 2 (2), 87–112.

Kerr, B., Godfrey-Smith, P., 2002. Individualist and multi-level perspectives on selection in structured populations. Biol. Philos. 17, 477–517.

Lehmann, L., Keller, L., West, S., Roze, D., 2007. Group selection and kin selection: two concepts but one process. Proc. Natl. Acad. Sci. 104 (16), 6736–6739.

Leigh, E.G., 2010. The group selection controversy. J. Evol. Biol. 23, 6–19.

Lion, S., Ajansen, V.A., Day, T., 2011. Evolution in structured populations: beyond the kin versus group debate. Trends Ecol. Evol. 26, 193–201.

Luo, S., 2013. A unifying framework reveals key properties of multilevel selection. J. Theor. Biol. 341, 41–52.

Marshall, J.A.R., 2011. Group selection and kin selection: formally equivalent approaches. Trends Ecol. Evol. 26, 325–332.

Matessi, C., Jayakar, S.D., 1976. Conditions for the evolution of altruism under Darwinian selection. Theor. Popul. Biol. 9, 360–387.

Maynard-Smith, J., 1964. Group selection and kin selection. Nature 201, 1145–1146.

Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.

Nowak, M.A., Tarnita, C.E., Antal, T., 2010a. Evolutionary dynamics in structured populations. Philos. Trans. R. Soc. 365, 19–30.

Nowak, M.A., Tarnita, C.E., Wilson, E.O., 2010b. The evolution of eusociality. Nature 466, 1057–1062.

Okasha, S., 2001. Why won't the group selection controversy go away? Br. J. Philos. Sci. 52, 25–50.

Okasha, S., 2006. Evolution and the Levels of Selection. Oxford University Press, Oxford.

Pinker, S., 2012. The false allure of group selection. Edge, June 21, 2012. ⟨http://edge.org/conversation/the-false-allure-of-group-selection⟩.

Price, G.R., 1970. Selection and covariance. Nature 277, 520–521.

Price, G.R., 1972. Extension of covariance selection mathematics. Ann. Hum. Genet. 35, 485–490.

Puhalskii, A., Simon, B., 2012. Discrete evolutionary birth-death processes and their large population limits. Stoch. Models 28 (3), 388–412.

Queller, D.C., 1992. Quantitative genetics, inclusive fitness, and group selection. Am. Nat. 139, 540–558.

Serway, R.A., Jewett, J.W., 2003. Physics for Scientists and Engineers. Brooks/Cole, Boston.

Simon, B., 2010. A dynamical model of two-level selection. Evol. Ecol. Res. 12, 555–588.

Simon, B., Fletcher, J.A., Doebeli, M., 2012. Hamilton's rule in multi-level selection models. J. Theor. Biol. 299, 55–63.

Simon, B., Fletcher, J.A., Doebeli, M., 2013. Towards a general theory of group selection. Evolution 67 (6), 1561–1572.

Simon, B., Nielsen, A., 2012. Numerical solutions and animations of group selection dynamics. Evol. Ecol. Res. 14, 757–768.

Sober, E., Wilson, D.S., 1998. Unto Others. Harvard University Press, Cambridge, MA

Szathmary, E., Demeter, L., 1987. Group selection of early replicators and the origin of life. J. Theor. Biol. 128, 463–486.

Traulsen, A., 2010. Mathematics of kin- and group-selection: formally equivalent? Evolution 64, 316–323.

Traulsen, A., Nowak, M.A., 2006. Evolution of cooperation by multilevel selection. Proc. Natl. Acad. Sci. USA 103, 10952–10955.

Traulsen, A., Sengupta, A.M., Nowak, M.A., 2005. Stochastic evolutionary dynamics on two levels. J. Theor. Biol. 235, 393–401.

Traulsen, A., Shoresh, N., Nowak, M.A., 2008. Analytical results for individual and group selection of any intensity. B Math Biol 70 (5), 1410–1424.

van Veelen, M., 2005. On the use of the Price equation. J. Theor. Biol. 237, 412–426.

van Veelen, M., 2009. Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. J. Theor. Biol. 259, 589–600.

van Veelen, M., 2011. A rule is not a rule if it changes from case to case (A reply to Marshal's comment). J. Theor. Biol. 270, 189–195.

van Veelen, M., Garcia, J., Sabelis, M.W., Egas, M., 2012. Group selection and inclusive fitness are not equivalent; the Price equation vs. models and statistics. J. Theor. Biol. 299, 64–80.

van Veelen, M., Luo, S., Simon, B. A simple model of group selection that cannot be analyzed with inclusive fitness, in review.

Wade, M.J., Wilson, D.S., Goodnight, C., et al., 2010. Multilevel and kin selection in a connected world. Nature 463, E8–E9.

West, S.A., Griffin, A.S., Gardner, A., 2008. Social semantics: how useful has group selection been? J. Evol. Biol. 20, 415–432.

Wild, G., Gardner, A., West, S.A., 2009. Adaptation and the evolution of parasitic virulence in a connected world. Nature 459, 983–986.

Wilson, D.S., 2012. The Central Question of Group Selection. Responses to Pinker (2012) From the Reality Club ⟨http://edge.org/conversation/the-false-allure-of-group-selection⟩.

Wilson, D.S., Wilson, E.O., 2007. Rethinking the theoretical foundations of sociobiology. Q. Rev. Biol. 82 (4), 327–348.

Wilson, E.O., 2013. The Social Conquest of Earth. Liveright Publishing, New York.