Adaptation by group selection: elimination acts on all levels

A Dissertation Presented by

Jeremy Khải Trần

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

December 2011

Stony Brook University

The Graduate School

Jeremy Khải Trần

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Lev Ginzburg – Dissertation Advisor
Professor, Ecology and Evolution

Stephan Munch - Chairperson of Defense
Associate Professor, School of Marine and Atmospheric Sciences

Daniel Dykhuizen – Committee Member
Distinguished Professor, Ecology and Evolution

John Damuth – University of California Santa Barbara,
Ecology, Evolution, and Marine Biology

This dissertation is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Dissertation

Adaptation by group selection: elimination acts on all levels

by

Jeremy Khải Trần

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

2011

Most evolutionary biologists today hesitate to invoke selection at any level higher than that of the individual, owing largely to the early criticisms against naïve group selection arguments. However, the result has been excessive caution in dealing with higher levels of selection, to the extent that it has been largely ignored, except by a handful of individuals. Yet the potential for work in this field remains great and has seen a recent resurge in interest.

The conclusions of this debate are not inconsequential. Group selection arose over the necessity to explain apparent adaptations of groups. However, there are three possibilities regarding the actual existence of group-level adaptation: groups are not adaptive, individual selection alone leads to adaptive groups, or higher-level selection leads to adaptive groups. The second possibility can be eliminated since we know that conflicts of interest can lead to individual

selection resulting in a "tragedy of the commons". Thus, either group traits are determined by individual selection regardless of their effect on group fitness or there exists a process that leads to groups possessing traits that are well-suited to overcoming obstacles to their persistence.

The appropriate unit for addressing selection depends upon the level of the question being asked. To address a question of group adaptation, we must use groups as our units. Individual-level approaches only track changes in the distribution of individuals; thus, they do little to answer group-level questions. Furthermore, individual-level approaches are often non-Markovian. The change in the distribution of individuals may also require knowing how those individuals are distributed into groups. One may be able to predict the future distribution of individuals, but cannot necessarily predict the future group distribution. On the other hand, a group-level approach tracks the change in the distribution of groups. When the group trait of interest is attributable to its composition, one can always determine the distribution of individuals by tracking the distribution of groups, while the reverse is not true.

Proceeding with groups as the unit, we can track the change of a group type by accounting for all inputs and outputs. In addition to new group formation and group extinction, we can account for group conversion to and from a group type to get a simple equation that is subsequently broken down further and rewritten in several forms to consider average fitness advantages between groups, individual fitness advantages within groups, migration, and drift. It is then applied to various scenarios to clarify and interpret evolutionary outcomes.

I develop a general framework for thinking about selection at various levels. Natural selection is the process by which configurations that are unable to persist are eliminated. As such, I generalize natural selection to be measured as proportional persistence (through survival or reproduction) in time, as opposed to number of replacements produced per lifetime, in order to resolve issues relating to multiple time scales. Persistence in nature is required at all levels of organization (as elimination at a higher level must also eliminate the constituent parts). This generally leads to individual and group interests coinciding. The traits (at any level of organization) of configurations that remain can be considered adaptive. Furthermore, constraints and tradeoffs that limit individually favored traits to those that result in stable higher levels can be considered adaptations.

The elimination of innately unstable configurations is an extreme form of higher selection equivalent to the elimination of inviable individuals. While this is an important process, it does not explain how natural selection can lead to solutions of higher group average fitness. However, higher level selection is not limited to this extreme. An innately stable configuration can be eliminated by competition with another innately stable configuration (*i.e.* both are viable, but one is inferior). I argue that this is the ideal case for illustrating the qualitative

importance of group selection, namely, when individual selection leads to multiple stable alternatives of different group fitness (*e.g.* asymmetric positive frequency-dependence). In contrast with other models of group selection, individual selection serves to maintain, rather than erode, variation between groups. Thus, favorable group traits can spread under a far wider range of conditions than with classic group selection. Yet, because individual selection can maintain either alternative, global fixation of the group advantageous trait is still attributable to group selection.

The issues addressed here are fundamental to the theory of evolution by natural selection. The idea that elimination processes can interact to shape patterns at multiple levels of biological organization has deep ramifications. The approach taken conceptually connects various approaches to the problem, while suggesting novel avenues for advancement.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my committee members Lev Ginzburg, Dan Dykhuisen, Steve Munch, and John Damuth for their feedback and support. I would also like to thank Charlie Janson, Massimo Pigliucci, Nick Friedenberg, Mark Lawley, and Christine Vendikos for their comments and discussion.

Chapter I: Introduction

1.1 Background

Most evolutionary biologists today hesitate to invoke selection at any level higher than that of the individual, owing largely to the early criticisms of Williams (1966) on this kind of thinking. However, the result has been excessive caution in dealing with higher levels of selection, to the extent that it has been largely ignored, except by a handful of individuals. Sober and Wilson (1998) describe the situation as follows:

> ...the rejection of group selection was celebrated as a scientific advance, comparable to the rejection of Lamarckism, that allowed biologists to close the book on one set of possibilities and concentrate their attention elsewhere. The memory of group selection had to be kept alive as an example of how not to think. It became almost mandatory for the authors of journal articles to assure their readers that group selection was not being invoked. A generation of graduate students learned to avoid group selection almost as if it were prohibited by one of the Ten Commandments.

Yet, the potential for work in this field remains great and has seen a recent resurge in interest (Nowak et al. 2010, Choi and Bowles 2007, Taylor and Nowak 2007, Hauert *et al.* 2007, Okasha 2006, and others).

Darwin (1871) was the first to suggest that selection could occur at the group level when he wrote:

> There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to give aid to each other and to sacrifice themselves for the common good,

> would be victorious over other tribes; and this would be natural selection.

However, the more recent debate over group selection began with Wynne-Edwards' (1962) argument that traits that increase the extinction risk of the group would be selected against at the group level, thus favoring helpful behaviors in spite of individual costs (which he used to attempt to explain population regulation).

Williams (1966) viewed such arguments invoking evolution for the good of the group as naïve, pointing out that such traits could not be favored since traits beneficial to the individual, regardless of their effect on the group, would quickly increase in frequency and spread to other groups faster than they could be eliminated. He and others further argued that group selection was not necessary to explain helpful behaviors since other mechanisms could explain them, such as kin selection (Hamilton 1964), reciprocity (Trivers 1971), game theory (Maynard Smith 1982), or other tradeoffs. Maynard Smith (1964) allowed that group selection as envisioned by Wynne-Edwards (1962) could occur, though only under a restrictive set of conditions.

1.2 Central questions

The conclusions of this debate are not inconsequential. Group selection arose over the necessity to explain apparent adaptations of groups. However, there are three possibilities regarding the actual existence of group-level adaptation: groups are not adaptive, individual selection alone leads to adaptive groups, or higher-level selection leads to adaptive groups. The second possibility can be eliminated since we know that conflicts of interest can lead to individual selection resulting in a "tragedy of the commons" (Hardin 1968). Thus, either group traits are determined by individual selection regardless of their effect on group fitness or there exists a process that leads to groups possessing traits that are well-suited to overcoming obstacles to their persistence. Hence, my goal is to simultaneously answer the following questions:

1.) Do group adaptations exist?

2.) Is anything gained from higher-level (*e.g.* group-level) thinking?

Demonstrating the former would provide a qualitative reason for adopting higher-level thinking. Yet, even if group adaptations do not exist, there may still be a quantitative reason for adopting higher-level thinking.

Okasha (2006) touches upon the major deficiency of past work when he states that “'real' group selection [...] must treat the groups as focal units, but this would imply that virtually none of the work purporting to be about group selection is actually about that subject.” Thus, my approach to addressing the central questions shall be to examine the consequences of using higher-level units to address the appropriate higher-level questions. In doing so, I shall demonstrate the theoretical importance of group selection, highlight the reasons that historic models of group selection have failed to become widely accepted, and suggest an ideal case for demonstrating group selection in nature.

Chapter II: Universal framework

2.1 Natural selection as a general process

The logic of natural selection applies to any entity that exhibits variation, differential persistence, and heritability. Variation in group types clearly exists, as every possible individual composition (the number of individuals of each individual type in a group) can be considered a different group type. Furthermore, groups may also differ in properties not attributable to their composition alone (*i.e.* truly emergent properties).

While it has traditionally been held that differential reproduction rather than persistence is the prerequisite, I have generalized to address the fact that persistence can be accomplished through either the persistence of self (survival) or the production of copies (reproduction). This is particularly important for groups, since they have no set lifespan and the average generation time of different group types are likely to be dissimilar. Therefore, offspring per lifetime or generation is not a good measure to compare rates of change. Instead, it is preferable to compare rates of persistence per standard unit of time, and hence explicitly incorporate survival as well as reproduction rather than implicitly, as with the traditional view. In doing so, the generalization stays true to the logic of natural selection while broadening its applicability.

In terms of groups, survival occurs when a group of a particular type continues to exist as that type, and reproduction occurs when a new group of a particular type is formed or when a preexisting group is converted to that particular type. Many aspects of group types can result in differential rates of persistence (*e.g.* groups composed of individuals with higher birth rates and/or lower death rates may go extinct less often and can fission more often as a result of faster growth).

Groups also exhibit heritability in that they tend to both survive as types similar to what they were previously and form new groups that are similar in composition to parent groups. It is worth noting that heritability need not be limited to Mendelian inheritance (or genetics for that matter), as the logic of natural selection holds in a more general sense. In fact, Darwin himself had little understanding of the mechanisms of heritability. While it is still the case that traits are passed on through time, the mechanism of heritability in groups tends to be different than is standard for individuals (and is in some sense Lamarckian).

In addition to individuals and groups, the logic of natural selection can apply to any level or organization that meets the three prerequisites and should occur simultaneously at all levels (Lewontin 1970).

2.2 Unit of selection

The appropriate unit for addressing selection depends upon the level of the question being asked. While an individual-level question might ask, “Why do individuals in a given group act cooperatively?”, a group-level question would ask, “Why are all groups composed of cooperative individuals?”. This distinction is not trivial.

One can model the dynamics of any unit as long as it is well-defined. The success of the model then depends upon proper accounting. Lower-level units often evolve within higher-level contexts. Dawkins (1976) suggested that an individual's genome or the composition of the group in which an allele is found can be considered part of that allele’s environment. However, to properly account for the changes in allele frequencies, one must take these "environmental" contexts into consideration (*e.g.* average fitness across contexts). This is not always simple since these contexts are not necessarily fixed and in fact can have their own evolutionary dynamics. For a unit to exhibit evolutionary dynamics, it must simply possess heritable variation with differential rates of persistence. Hence, it becomes necessary to determine higher-level dynamics in order to determine the dynamics at a lower level, except in certain reducible cases.

Individual-level "group selection" approaches (MLS1 following Damuth and Heisler 1988, Okasha 2006) track changes in the distribution of individuals while allowing individual fitness to depend on the group context (*e.g.* frequency-dependence). When the group type is determined by its individual composition (the number of each individual type it contains), the individual distribution, I, is given by the number of total individuals that are of a given type for each type:

$$I_X = \sum_i P_i f_{X_i} \qquad (1)$$

for each individual type *X*, where P_i is the proportion of groups of type *i*, and f_{X_i} is the number of individuals of type *X* in a group of type *i*. Since the fitness of an individual depends on its context, a type's average fitness is given by

$$\bar{w}_X = \sum_i p_{X_i} w_{X_i} = \frac{\sum_i P_i f_{X_i} w_{X_i}}{\sum_i P_i f_{X_i}} \qquad (2)$$

where p_{X_i} is the proportion of individuals of type *X* found in groups of type *i*, and w_{X_i} is the fitness of an individual of type *X* in a group of type *i*.

A problem with these individual-level approaches is that they often are non-Markovian (Glymour 2008). Assuming heritability, the average fitness of each type will influence the change in the individual distribution from one generation to the next. Since the group distribution is used to determine the average fitness (Eq. 2), determining the change in the distribution of individuals may require knowing how those individuals are distributed into groups. Even given the current distribution of groups, one may be able to predict the future distribution of individuals for one time step but cannot necessarily determine the future group distribution needed to calculate for further time steps. This is a serious limitation when determining ultimate evolutionary outcomes, which discussion of adaptation necessitates.

This issue is avoided in cases that are reducible (in a quantitative sense) to standard individual-level selection (independent of group context), such as when group context does not affect individual fitness. In that case, w_{X_i} is a constant, w_X, for all *i*, and therefore, $\bar{w}_X = w_X$. MLS1 approaches are also reducible to standard individual-level selection when the group distribution *G* (knowing P_i for all *i*), can be determined from the individual distribution, *I* (knowing p_X for all *X*). This may be the case when groups reform each generation as with Wilson's (1975) intrademic model. This is because if $I_{t+1} = f(I_t, G_t)$ and $G_t = f(I_t)$, then $I_{t+1} = f(I_t)$, where *f* is not a specific function but rather *y=f()* denotes that *y* is some function of the variable(s) in parentheses. It is then also possible to determine *G* at any time from *I*, and can therefore address group-level questions, though the validity of groups as evolutionary units in this case is questionable. Thus, MSL1 cases are either reducible, non-Markovian, or become MLS2 (by using $G_{t+1} = f(G_t)$).

On the other hand, group-level approaches (MLS2 following Damuth and Heisler 1988, Okasha 2006) track the change in the distribution of groups. When the group trait of interest is attributable to its composition, one can always determine the distribution of individuals by tracking the distribution of groups, while the reverse is not true. That is to say that $I_t = f(G_t)$, but, except in special cases, $G_t \neq f(I_t)$. That is because for each individual type *X*, Eq. 1 allows for only one solution of p_X given a group distribution, while there are many possible group distributions that would satisfy the equation for a given p_X. Thus, MLS2 approaches can address both individual- and group-level questions whenever group type is determined by its composition. Hence, for individual-level questions, MLS2 approaches are capable of, though not required for, addressing reducible cases ($I_{t+1} = f(I_t)$) (Table 1). At the same time, they are capable of addressing group-level questions where individual-level approaches cannot (*i.e.* when $G_t \neq f(I_t)$ and when group traits are not attributable to group composition).

It should be noted that while the Price equation (Price 1972) can use groups as units, it typically remains non-Markovian. This is because it only determines the change in average characteristic, therefore losing the distributional information necessary for subsequent time steps. However, more general approaches that track group distributions do not have this problem.

Hence, while MLS1 approaches acknowledge that fitness depends on context, MLS2 approaches acknowledge that these contexts may themselves be evolving entities. This additional level forms the basis for the distinction between individual-level and group-level questions. Therefore, to address questions of group adaptation, we must use groups as our units and are thus taking a MLS2 approach.

Table 1. Domains of applicability of different approaches to tracking evolutionary change.

Approach	Addresses individual-level questions	Addresses group-level questions	Markovian
Individual $I_{t+1} = f \;\; I_t$	Yes	Only when $\mathrm{G}_t = f \;\; I_t \;\;,$ then $G_{t+1} = f \;\; I_t$ Never when group type is not attributable to composition	Yes
MLS1 $I_{t+1} = f \;\; I_t, G_t$	Yes	Only when $\mathrm{G}_t = f \;\; I_t \;\;,$ then $G_{t+1} = f \;\; I_t$ Never when group type is not attributable to composition	Only when $\mathrm{G}_t = f \;\; I_t$, then $I_{t+1} = f \;\; I_t$ and is reducible to an individual approach Or when $G_{t+1} = f \;\; G_t$, then $I_{t+1} = f \;\; G_t$ and becomes a MLS2 approach
MLS2 $G_{t+1} = f \;\; G_t$	Whenever group type is attributable to composition, then $I_{t+1} = f \;\; G_{t+1}$ and $I_{t+1} = f \;\; G_t$	Yes	Yes

2.3 A general equation to track groups

Proceeding with groups as our units, one must determine how to account for all of the inputs and outputs from a group type. Inputs include the formation of new groups of a particular type and the conversion (*e.g.* a change in composition resulting from individual selection within groups or migration from other groups) to a particular group type from another. Outputs include extinction and the conversion to other group types. Hence, for the number of a particular group type, N, and time given by t

$$\frac{\Delta N}{\Delta t} = \frac{N_{formed} - N_{extinct}}{\Delta t} + \frac{N_{convert\,in} - N_{convert\,out}}{\Delta t} \tag{3}$$

Using Eq. 3, we can determine the conditions, assuming only two group types, under which one group type (A) will increase in proportion relative to another (B), by comparing their *per capita* rates of change and rearranging to get

$$\begin{gathered}\left[\left(\frac{A_{formed}}{A_{initial}\cdot\Delta t} - \frac{A_{extinct}}{A_{initial}\cdot\Delta t}\right) - \left(\frac{B_{formed}}{B_{initial}\cdot\Delta t} - \frac{B_{extinct}}{B_{initial}\cdot\Delta t}\right)\right] + \\ \left[\left(\frac{B_{convertA}}{A_{initial}\cdot\Delta t} - \frac{A_{convertB}}{A_{initial}\cdot\Delta t}\right) - \left(\frac{A_{convertB}}{B_{initial}\cdot\Delta t} - \frac{B_{convertA}}{B_{initial}\cdot\Delta t}\right)\right] > 0\end{gathered} \tag{4}$$

Though clearly an oversimplification when intermediate group types exist, categorizing groups as either A-type or B-type allows us to apply Eq. 4 to a variety of historical models in a way that allows for clear understanding of the evolutionary outcomes within a group-level framework.

Chapter III: Historic group selection models

3.1 Shortcomings of historic models

There are two main shortcomings of historic conclusions regarding group selection. The first is, by focusing on altruism, individual and "group selection" have been placed as opposing rather than simultaneous hierarchical processes. The erroneous view was that altruism would prevail if "group selection" were stronger than individual selection, with the assumption that "group selection" favors altruists and individual selection favors cheaters. However, it shall be seen that group selection generally favors group types that are stable (with respect to individual selection), which groups of altruists susceptible to invasion by cheaters are not. The second is the dismissal of group selection whenever an individual advantage can be found for an individual trait that benefits other members of the group. However, it shall be seen that group selection generally leads to coinciding individual and group interests.

3.2 Classic group selection

Classic group selection involves a trait that is beneficial to the group at a cost to the individual. Thus, there are two types of individuals: altruists (A) and cheaters (B) (Fig. 1). There are two equilibrial group types: groups fixed for altruists and groups fixed for cheaters. The former is an unstable equilibrium since, if any cheaters are introduced, cheaters will increase in frequency as a result of individual selection. The latter is a stable equilibrium since any altruists that arise will be outcompeted by cheaters and therefore eliminated by individual selection.

For simplicity, let us consider just the group types: A-dominant and B-dominant. While this coarse grouping is not very precise (more precise grouping will be addressed later), it provides a reasonable approximation, particularly when mixed groups are transitory between the two equilibrial types. We can then use Eq. 4 to assess the conditions under which altruism can prevail.

The first term (bracketed) is positive since group formation and/or survival rates are presumably higher in altruist groups than in cheater groups (at least in the case when there is a group advantage, meaning that the average individual fitness is higher). However, the second term tends to be negative since cheaters can readily invade groups of altruists, while the reverse is not true. Therefore, the outcome depends on the relative magnitude of these terms.

The first term is maximized when there is a large group benefit from altruism. The magnitude of the second term is minimized when: 1) introduction of cheaters into groups of altruists is rare (*e.g.* low gene flow and mutation rates), 2) introduction of cheaters is unlikely to result in their fixation (*e.g.* low cost of altruism and high genetic drift), and 3) cheater groups are frequently converted to altruist groups (*e.g.* low cost of altruism and high genetic drift). Thus, high group benefit, low cost of altruism, low migration rates, low mutation rates, small population sizes, high variability in fitness, and frequent founder effects tend to favor the spread of altruistic groups. Others have previously converged upon similar conditions (Maynard Smtih 1964, Levins 1970, Levin and Kilmer 1974, Gilpin 1975, Boorman and Levitt 1980). However, stochastic processes in particular render the conditions under which altruism can evolve more robust than many early models allowed (*e.g.* those that assumed that a single introduction of a cheater would necessarily lead to its fixation). Thus, while echoing the conclusions of historical work, this framework provides a fairly simple and intuitive approach that, as we shall later see, can also be applied more broadly.

A naïve group selection argument, which assumes that groups with the highest average individual fitness prevail, would conclude that altruism would prevail over cheating because average individual fitness is higher in groups of altruists than in groups of cheaters. A naïve individual selection argument, which assumes that all groups end up at the stable equilibrium due to individual selection within groups, would conclude that cheaters would prevail because they have higher fitness than altruists in all contexts in which they coexist and can invade pure groups of altruists. However, we see that either type can prevail, depending upon the conditions. This constitutes a qualitative reason for incorporating a group-level approach. Yet, a lack of compelling cases of classic group selection in nature has led skeptics to dismiss the importance of group selection altogether. There are in fact inherent shortcomings of classic group selection that make it less than ideal to demonstrate in nature, but this does not discount the process.

The classical argument has been that the conditions necessary are rare (Maynard Smith 1964). However, while it is clear that Eq. 4 needs to be satisfied for the spread of groups of altruists (and therefore altruistic individuals), I would argue that the required set of conditions need not be rare. It is not necessary for each of the conditions discussed above to be met in the extreme. For example, if the cost of altruism is minor, such that it is nearly neutral, only a moderate group benefit and minor stochastic effects would be sufficient. Even when the conditions for the spread of altruism are not met at the group level, they are likely to be met at the species level (due to low gene flow between species). However, at this level of organization, the rate of change in species type involves mutation rates (and subsequent fixation) and diversification rates of species and, consequently, what constitutes an altruistic act may be quite different.

Despite this, there is a difficulty in that, as with any study on natural selection, to observe the process, variation must currently exist. Hence, if an inferior alternative has already been eliminated, there no longer exists an alternative to which one can compare fitness. When variation exists, either it is rare (*e.g.* in selection-mutation balance), fortuitous (*e.g.* still in the process of one being eliminated, different selection pressures in different regions, etc.), or not classic group selection (*e.g.* experiencing negative frequency-dependent selection). In essence, the more effective the process, the more difficult it is to observe it. In such a case, one should not necessarily presume rarity of occurrence from scarcity of evidence.

The problem with demonstrating classic group selection is not just in finding observable cases in which the conditions are met, but also in distinguishing from alternative arguments when they are met. The main difficulty lies in demonstrating a true cost of a group-beneficial behavior. Critics will often argue that there is an individual benefit to counteract the cost that was not considered. Adamant individual-selectionists, assuming that there must be an individual benefit, may even argue this point without providing any specific suggestion of how individuals may benefit. Thus, group-selectionists may be left with the impossible task of disproving a negative.

Kin selection in particular is often cited to explain apparent altruism that can be discounted, yet it is incredibly difficult to measure and correct for. To do so would require precise fitness measures in addition to complete knowledge of relatedness. Furthermore, this problem is almost impossible to avoid because the conditions that allow classic group selection to favor altruism tend to be the same conditions that increase association with kin (*e.g.* small population sizes, low gene flow, founder effects, etc.). However, it is important to keep in mind that classic group selection does not require group members to be related (*i.e.* altruism can be favored in the absence of any kin selection) and that relatedness to other group members need not be constant. It is also worth noting that the entire concept of inclusive fitness has recently come under attack on grounds that standard natural selection theory is a simpler and superior approach (Nowak et al. 2010).

While some have accepted natural cases, such as the evolution of avirulence in the myxoma virus (Sober and Wilson 1998) or biased sex ratios in *Anelosimus eximius* (Aviles 1993), issues of relatedness continue to lead others to question them. Even leaf sharing to unrelated individuals in *Acromyrmex versicolor* (Rissing *et al.* 1989) may in fact result from a form of punishment (Pollock *et al.* 2004). Thus, while classic group selection has the theoretical potential to provide a qualitative reason for considering group selection, it is extremely difficult to provide a convincing case in practice.

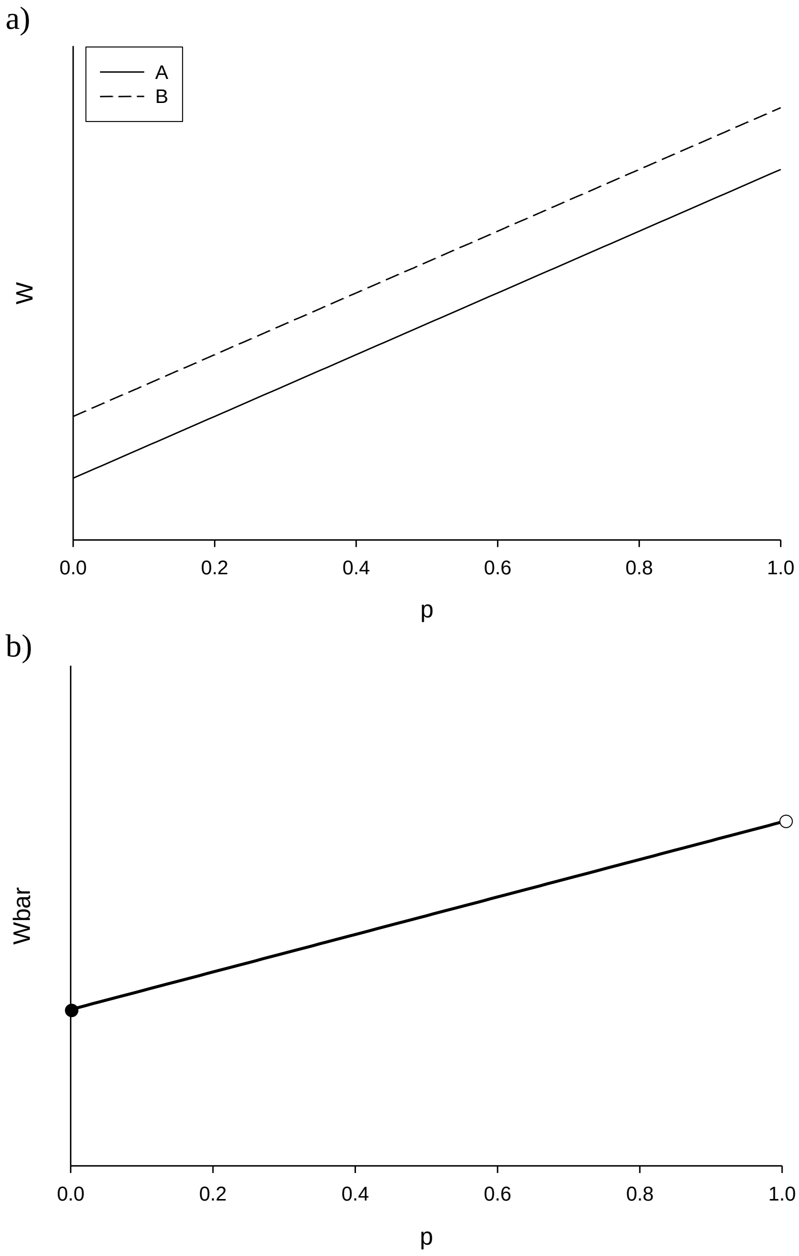


Figure 1. Classic group selection. (a) Individual fitness as a function of p, the proportion of individuals of type A (altruists). (b) Average fitness as a function of the proportion of individuals of type A, calculated as $w_A \times p + w_B \times (1-p)$. Solid points are stable equilibria; hollow points are unstable equilibria.

3.3 Intrademic selection

Wilson (1975) presented a different approach to the evolution of helpful (not truly altruistic) traits in his model of intrademic selection. In Wilson’s model, individuals live in what he calls trait groups, subpopulations in which ecological interactions occur, a number of which form a randomly-intermating deme. At the end of a period of selection, there is a dispersal/mixing phase each generation before individuals form new trait groups. In this scenario, a trait could decrease within trait groups, but still increase within the entire deme, as long as variation between trait groups is sufficient. This occurs because trait groups with a high frequency of helpers result in high reproduction of both helpers and non-helpers in that trait group, but helpers are more likely to be in groups with high frequency of helpers as a result of types being non-randomly distributed. Conversely, trait groups with a low frequency of helpers result in low reproduction of both types within that trait group, but non-helpers are more likely to occur in this situation. Essentially, intrademic selection allows for helpful traits to be directed preferentially towards other helpers. As such, being a helper can be viewed as being individually beneficial when helpers receive on average more benefit than non-helpers, and the difference in benefits received (determined by the amount of variance) is greater than the cost.

A parallel can be drawn to altruism in that helping incurs a cost; however, this is offset by the benefit derived from preferentially associating with other helpers. While it is also the case that altruists are more likely to occur in groups with other altruists in scenarios in which altruism prevails, the association need only be correlational. With helpers under intrademic selection, the association is causal. The key difference is the mechanism by which variance is maintained. It is different from classic group selection in that the group that the individual is part of is transitory and membership is influenced by the individual's trait. As such, intrademic selection is reducible to individual selection since the distribution of trait groups can be determined solely from the previous distribution of individuals (one does not need to know the previous distribution of trait groups). This prevents it from being useful in demonstrating a qualitative importance of group selection.

However, Wilson's (1975) work made an important contribution by showing how, at the individual level, non-homogenous interactions can favor traits with apparent costs. This argument of non-homogenous interactions is more general than the specific model above involving trait groups. In fact, it can apply to classic group selection and other cases. However, being a MLS1 approach, it is limited by being non-Markovian.

The non-homogenous interactions argument is also more general than kin selection. Associating with kin is simply a means by which variance between groups can be increased because kin are likely to be of the same type. Kin selection is often viewed as working through a direct cost with indirect benefits (through kin). However, one can also view it as working through direct benefits gained from kin, with direct or indirect costs, completely analogous to the individualistic view of intrademic selection presented above. Goodnight *et al.* (1992), through contextual analysis (an analysis of covariance; see Heisler and Damuth 1987) come to the same conclusion that the evolution of helpful traits depends on high variation between groups and find an equivalency to Hamilton's rule (1964) when expanded to consider correlation of individuals of the same type rather than relatedness. Thus, any type of assortative interaction has the same result, whether it is by kin association or some ability to recognize and associate with other helpers. Nevertheless, the limitations of a MLS1 approach still apply.

Let us then consider reducible intrademic selection within demes, but allow for selection between demes. When assortivity is low, the situation is essentially that of classic group selection (Fig. 1). However, when assortivity increases to the point that the benefits helpers receive from other helpers exceeds the cost of helping, it switches to a scenario in which helping is both beneficial to the individual and the group (Fig. 2). The stable equilibrium is at fixation for helpers. Therefore, groups of helpers not only have an advantage in terms of group formation and extinction, but also in terms of group conversion (stability).

Both a naïve individual selection and a naïve group selection argument would correctly predict that helpers would prevail.

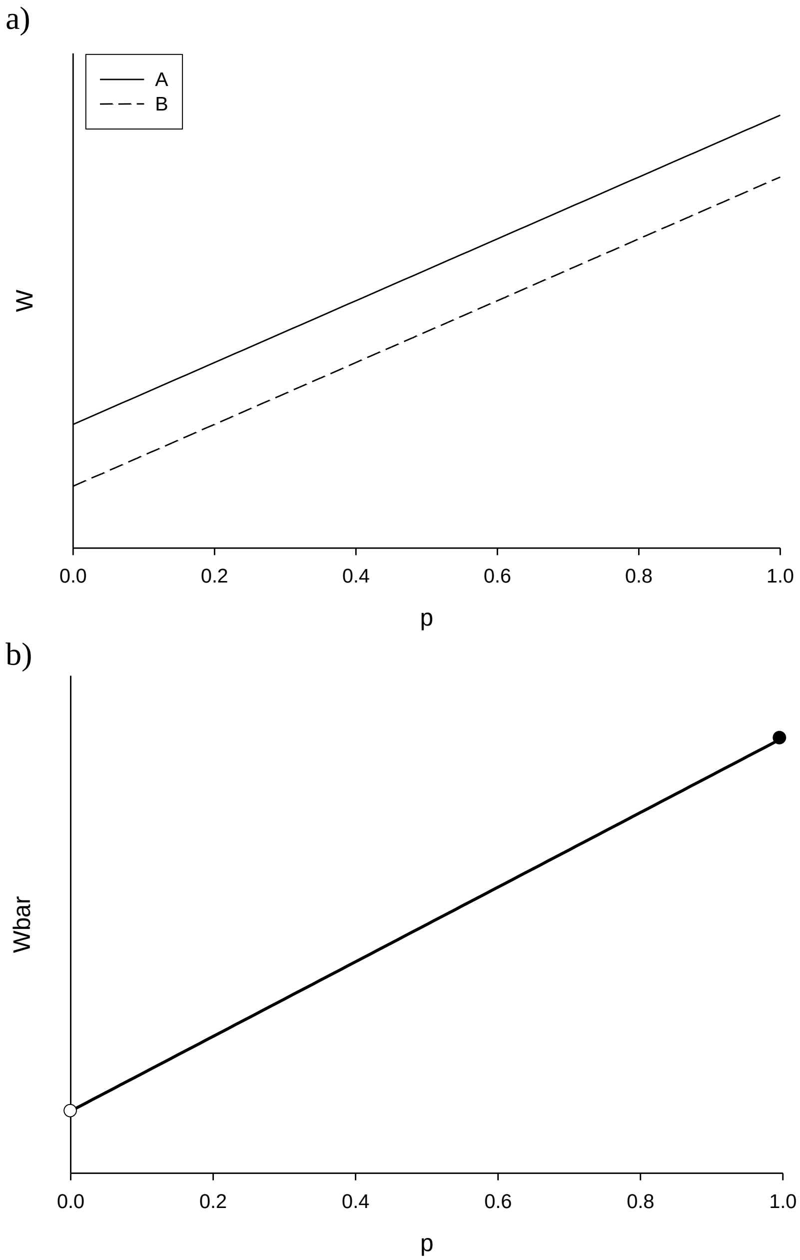


Figure 2. Interdemic selection with variance several times higher than random binomial variance; equivalent to a group beneficial trait with an individual advantage. (a) Individual fitness as a function of p, the proportion of individuals of type A (helpers). (b) Average fitness as a function of the proportion of individuals of type A, calculated as $w_A \times p + w_B \times (1 - p)$. Solid points are stable equilibria; hollow points are unstable equilibria.

Chapter IV: Identifying a more ideal case to demonstrate group selection

4.1 Alternative ways in which individual and group interests interact

We have now seen a scenario in which a group beneficial behavior is individually disfavored and one in which it is favored, with neither being ideal for demonstrating a qualitative reason for considering group selection. However, there are other ways in which individual traits with group consequences can interact. For example, they can be neutral, exhibit negative frequency-dependence, or exhibit positive frequency-dependence.

If two individual traits are neutral relative to each other at the individual level, but one has a positive group effect, a naïve individual-level approach would predict that natural selection would not favor either trait and that the evolutionary outcome would depend on drift. However, from Eq. 4, we can see that group formation and extinction would favor groups with more helpers while conversion rates would be essentially neutral. In this instance, a naïve group selection would correctly predict the deterministic increase of the neutral group-beneficial trait. Yet, this result is unsurprising, and probably accepted by even the most adamant individual selectionist, by allowing that group beneficial traits can prevail in the absence of individual selection. However, the domain of applicability is a highly restricted and, experimentally, it can be near impossible to ascertain whether a trait is neutral or weakly favored (and if weakly disfavored, it would fall under classic group selection).

With negative frequency-dependence, in which rare types are favored, average individual fitness is highest in mixed groups (Fig. 3). While both a naïve individual selection and naïve group selection approach would predict coexistence, it should be noted that the highest average fitness need not be at the stable equilibrium (Roughgarden 1998). Such a scenario may be interesting from a quantitative perspective to determine the equilibrial global distribution but is not particularly informative when seeking qualitative reasons for incorporating group selection.

Positive frequency-dependence, in which rare types are disfavored, is perhaps the most interesting case (particularly asymmetric positive frequency-dependence) for demonstrating a qualitative importance of group selection. This results in two stable equilibria, which are evolutionarily stable strategies (ESS; see Maynard Smith 1982), at fixation for either trait in addition to an unstable threshold, which is a Nash equilibrium (in which no individual can benefit by unilaterally changing strategy; Nash 1950) but not an ESS (Fig. 4). A naïve group selectionist would predict that the trait that has higher average fitness when fixed would prevail. How a naïve individual selectionist would predict the evolutionary outcome is less clear, since there is more than one stable group type. If one ignored the group consequences, one would perhaps argue that, with low gene flow, each group would become fixed for one individual type or the other, depending on which was initially more common, and persist in that state, and that with a more homogenous population, the initially more abundant trait would prevail globally. In any case, neither naïve approach satisfactorily predicts the evolutionary outcome in all cases since the individual type that prevails depends on which group context prevails (a group-level question).

We can see from Eq. 4 that, like the other cases, the group type that prevails is the one with the best combination of high average fitness (which implies high formation rates and/or low extinction rates) and stability (which relates to the conversion rates as influenced by gene flow and individual selection). The conditions necessary to favor the group type with higher average fitness are far more robust than with classic group selection since individual selection against rare types minimizes the success of invasions by the other type. We can also see, from a MLS1 perspective, that the high variance between groups favoring the group beneficial trait is maintained by positive frequency-dependent selection rather than stochastic processes, as with classic group selection, or assortative interactions, as with intrademic selection.

The potential of asymmetric positive frequency-dependence for favoring group beneficial traits has been recognized (Boyd and Richerson 1990), but its impact on the group selection debate has been minimal. This is probably largely due to the paradigm of using altruism as the illustrative case. However, asymmetric positive frequency-dependence is the most useful and practical demonstration of the qualitative importance of group selection. Individual selection can lead to multiple ESS, but it is group selection that is the process that generally leads to evolution favoring the ESS that is also a strong Nash equilibrium (meaning the Nash equilibrium in which there is no coalition of individuals that could benefit by changing strategy; *i.e.* the ESS with the highest average fitness). It should be cautioned, however, that group selection can also favor a lower ESS if it is more stable than the alternatives (and enough so to compensate for lower average fitness). Thus, the location of the threshold (unstable equilibrium), in addition to the factors discussed with classic group selection, is important. Yet the utility of a MLS2 approach is in its ability to correctly determine the evolutionary outcome in either case.

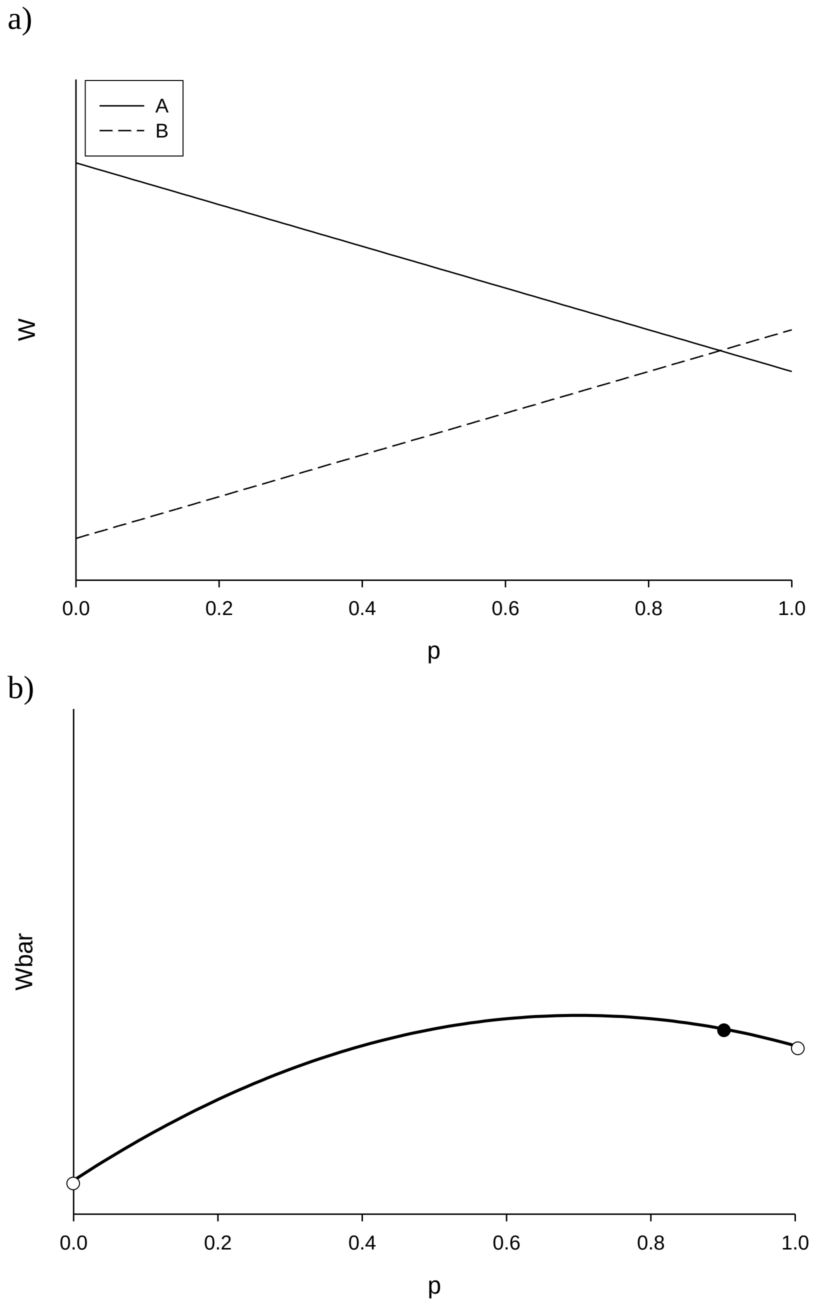


Figure 3. Negative frequency-dependent selection. (a) Individual fitness as a function of the p, proportion of individuals of type A. (b)Average (group) fitness as a function of the proportion of individuals of type A, calculated as $w_A \times p + w_B \times (1-p)$. Solid points are stable equilibria; hollow points are unstable equilibria.

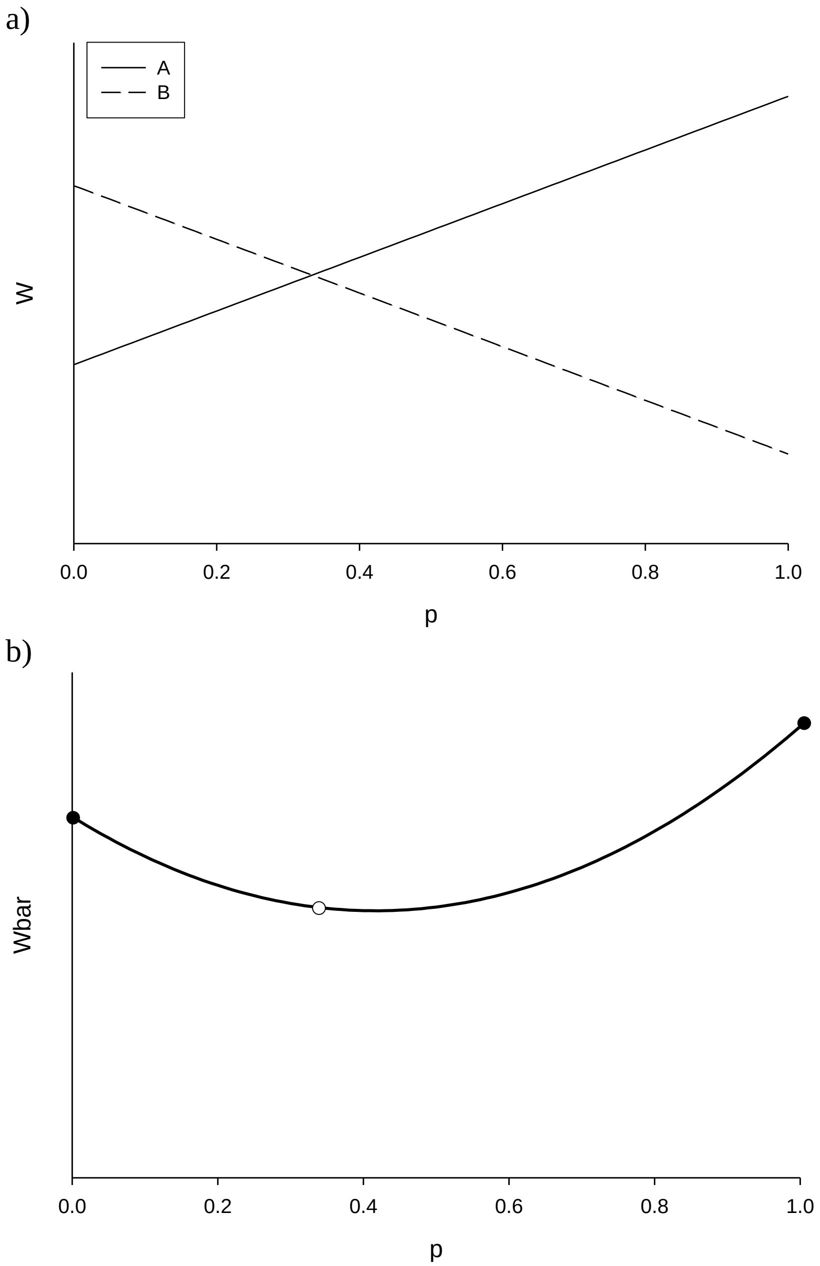


Figure 4. Positive frequency-dependent selection. (a) Individual fitness as a function of p, the proportion of individuals of type A. (b)Average (group) fitness as a function of the proportion of individuals of type A, calculated as $w_A \times p + w_B \times (1 - p)$. Solid points are stable equilibria; hollow points are unstable equilibria.

4.2 Potential asymmetric positive frequency-dependence in nature

Asymmetric positive frequency-dependence occurs when rare types are individually disfavored (and thus conformity is favored), but groups conforming to one type can outperform groups conforming to another. There are numerous natural examples where we know, or at least suspect, this to be case. These include aposematic coloration, color patterns to blend in with a herd, reproductive timing, and various forms of reward and/or punishment.

With aposematic coloration, conformers benefit through signaling unpalatability to potential predators. The cost incurred through mortality resulting from predator learning is diluted among similarly colored aposematic individuals. Hence, the average predation risk of an aposematic individual in a group of other aposematic individuals may be quite low, even more so than a cryptic individual. However, being the only aposematic individual carries a high risk since being conspicuous without predators recognizing the signal makes a lethal encounter likely. Thus, in a group of cryptic individuals, it is better to also be cryptic. Therefore, group selection may be necessary to explain how aposematism spreads once it arises. Furthermore, it suggests a process in which better aposematic signals can replace less effective ones. Especially in this case, kin selection is not a viable alternative since individuals benefit from kin that exhibit the same signal as themselves regardless of which signal the individual exhibits (*i.e.* it does not determine which aposematic signal prevails).

Conformity of colors and patterns in herds or schools is a very similar situation, except that conformity is favored because individuals that stand out are more easily isolated by predators. Therefore, conformity is like camouflage to blend in amongst other individuals. Again, conformity for particular patterns, such as the disruptive patterns of zebras, may be more effective than conformity for other patterns. As such, group selection can lead to more effective patterning.

Group selection may also be very important in reproductive or other timing, since the optimal time for reproductive effort, whether that be releasing of pollen or performing a mating display, depends not only on climatic seasonality, but also on the timing of other individuals. Hence, group selection can potentially lead to more effective timings by eliminating those that are suboptimal.

Even reciprocity, often used as an alternative explanation for group selection, is actually a positive frequency-dependent trait. An initiator of a helpful behavior would suffer a cost that it would not recuperate if there are no other individuals willing to reciprocate. However, non-reciprocators do not do well in groups of reciprocators once discovered and benefits are withheld (which can be viewed as a type of punishment).

Rewards and punishments can actually be used to reconcile individual and group interests. Enforcement, when there is a cost involved, may require that individuals enforce not only for the trait of interest, but also against non-enforcers, in order to be stable (Maynard Smith 1999). Enforcing then becomes positive frequency-dependent. When there are few enforcers, the cost of enforcing is too high relative to the risk of being caught for not enforcing. When there are many enforcers, the cost of enforcing is relatively low compared to the risk of being caught for not enforcing. Enforcement can then favor or disfavor any arbitrary trait.

Dugatkin and Godin (1992) showed that, in guppies, female preference is influenced by the preference of other females. This conformity is in line with the "sexy son hypothesis" (Weatherhead and Robertson 1979), which posits that to maximize fitness, females should choose males that other females find attractive, so that the male offspring will possess the desired traits. However, which traits are desired can be arbitrary. Interestingly, in the case of the guppies, this conformity is socially rather than genetically driven (Dugatkin and Godin 1992).

However, group selection has the potential to remove the arbitrariness by opposing enforcement for group detrimental traits and promoting enforcement for group beneficial traits (Fig. 5). Thus, certain traits can reconcile individual and group interests by making it individually beneficial to behave in a way that is beneficial to the group. These reconciling traits should prevail evolutionarily, since they are still stable, but have higher average group fitness. Hence, natural selection favors groups in which conflicts of interest are resolved. This emphasizes the point that individual-level explanations do not preclude a higher-level process. Just as single dimension approaches are reminiscent of Wright's shifting balance, multidimensional approaches are reminiscent of adaptive landscapes (Wright 1980); however, here we are explicitly taking into account individual and group interests in determining stable points on a surface that do not necessarily coincide with average fitness peaks.

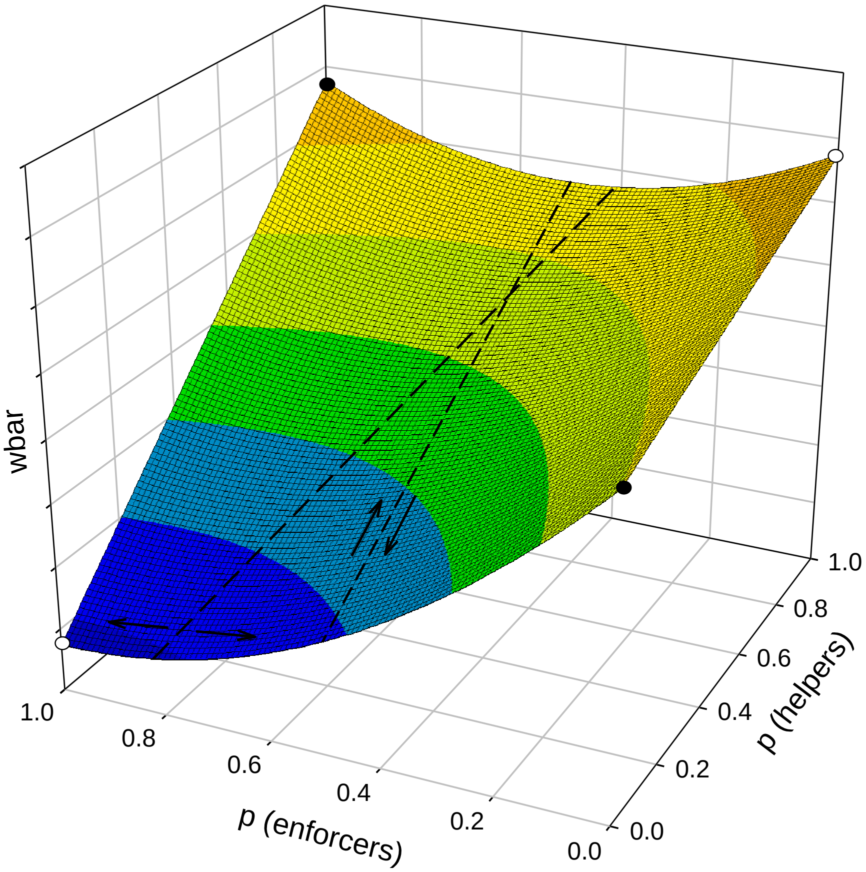


Figure 5. Average fitness as a function of the proportion of individuals that are helpers and the proportion that are enforcers. Solid points are stable equilibria; hollow points are unstable equilibria. Dashed lines are unstable equilibrium lines, with arrows indicating direction of change. Average fitness was calculated as follows:

$$w_{helper/enforcer} \times p_{helper/enforcer} + w_{helper/non-enforcer} \times p_{helper/non-enforcer} + w_{non-helper/enforcer} \times p_{non-helper/enforcer} + w_{non-helper/non-enforcer} \times p_{non-helper/non-enforcer} = (a + bp_h - c_e(1-p_e) - c_e(1-p_h) - c_h)(p_e p_h) + (a + bp_h - c_p p_e - c_h)((1-p_e)p_h) + (a + bp_h - c_e(1-p_e) - c_e(1-p_h) - c_p p_e)(p_e(1-p_h)) + (a + bp_h - 2c_p p_e)((1-p_e)(1-p_h))$$

Where w = fitness and p = proportion of individuals that are of the type designated in the subscript, c_h = cost of helping n individuals, c_e = cost of enforcing against n individuals, c_p = cost of being “punished” by n individuals, b = benefit of being “helped” by n individuals, n = number of individuals in the deme, p_e = proportion of individuals that are enforcers, p_h = proportion of individuals that are helpers, and a = fitness in the absence of any costs or benefits. Shown is with $a = 1$, $c_h = 0.25$, $c_e = 0.35$, $c_p = 0.5$, and $b = 0.6$.

Chapter V: Relaxing assumptions

5.1 Fine group-typing

The coarse group-typing used thus far is conceptually useful, though clearly a mathematical oversimplification. To be more precise, one needs to take intermediate frequencies into account as well. Additionally, group size is an important group attribute that potentially effects individual fitness and migration rates (*e.g.* through density-dependence) and extinction/formation rates (*e.g.* smaller group are more likely to go extinct and larger groups are more likely to split to form new groups).

Let us then consider an example in which a group type is precisely defined by its individual composition in terms of number of individuals of each type. There are six processes that can lead to a transition from one group type to another. Groups can gain individuals of any type through birth or immigration, lose individuals through death or emigration, fission into multiple smaller groups, or fuse with other groups. For our example, we shall simplify by making it a continuous model, such that transitions due to birth, death, and migration are non-zero only to adjacent group types (plus or minus one individual of any type) and that fission only results in splitting into two groups (Fig. 6). We shall also ignore group fusion, though incorporating it is fairly straightforward. So that all of the group types are easily organized on a two-dimensional plane, our example will be a haploid model with only two individual types (a diploid model with two alleles would have 3 individual types). To make the number of group types finite, we shall use a maximum group size determined by a ceiling carrying capacity (in this case, 10 individuals), and only groups smaller than that shall receive immigrants. Birth rates are constant for all group sizes smaller than the carrying capacity and are equal for both individual types. Mutation rates are symmetric. Emigration rates are equal for all individuals. Only groups of size 10 fission and do so according to a binomial distribution with probability 0.5 of sorting to either new group. We shall use a case of classic group selection for our first example by making death rates higher for altruists (A-types) and death rates for both types decreasing as the proportion of other group members being altruists increases. Our second example shall be an asymmetric positive frequency-dependent model with the same average fitness for pure groups as the first example, but with a threshold where the proportions of the two individual types are equal.

This information is then used to construct a transition matrix between group types (with transitions as seen in Fig. 6). It should be noted that the transition matrix changes with the group distribution because for any given group type, the immigration rate for each individual type depends on how many of each individual type are emigrating from other groups. Many of the assumptions made above could be altered, but as long as transitions between group types can be defined, the dynamics can be modeled. Depending on the complexity of the model, it may be feasible to solve for equilibrial group distributions by solving for group distribution vectors that when multiplied by the transition matrix (which may be a function of that vector) results in a vector of zeros and to determine their stability by examining the matrix of partial derivatives. If not feasible, numerical simulation may be appropriate. However, as the number of group types increases (*e.g.* more individual types, larger maximum group size, etc.), computational limits are quickly reached. Nonetheless, it is conceptually useful to understand transitions between group types and their evolutionary consequences.

For the classic group selection example, we can see that with low migration rate, groups of altruists quickly increase due to their higher growth rates and subsequent fissioning. However, with higher migration rates, we observe a higher occurrence of mixed group types, which tend to transition to cheater groups (Fig. 7). We can also see that at an intermediate migration rate, the initial distribution influences the evolutionary outcome (Fig. 8) as does the maximum population size (Fig. 9). However, the evolutionary outcome for the asymmetric positive frequency-dependent case is far more robust to these factors (Fig. 10-12).

It is also possible to extend this deterministic model, which assumes an infinite number of groups, to a stochastic model. This would incorporate group-level drift. To do so would require extending the state space to include all possible combinations of group types up to the maximum number of groups. The transitions between states in a continuous framework (in which only one group can transition at a time) are given by the transition rates of the deterministic model corresponding to the stochastic transition. However, since the number of states can be quite large, simulation may be preferable.

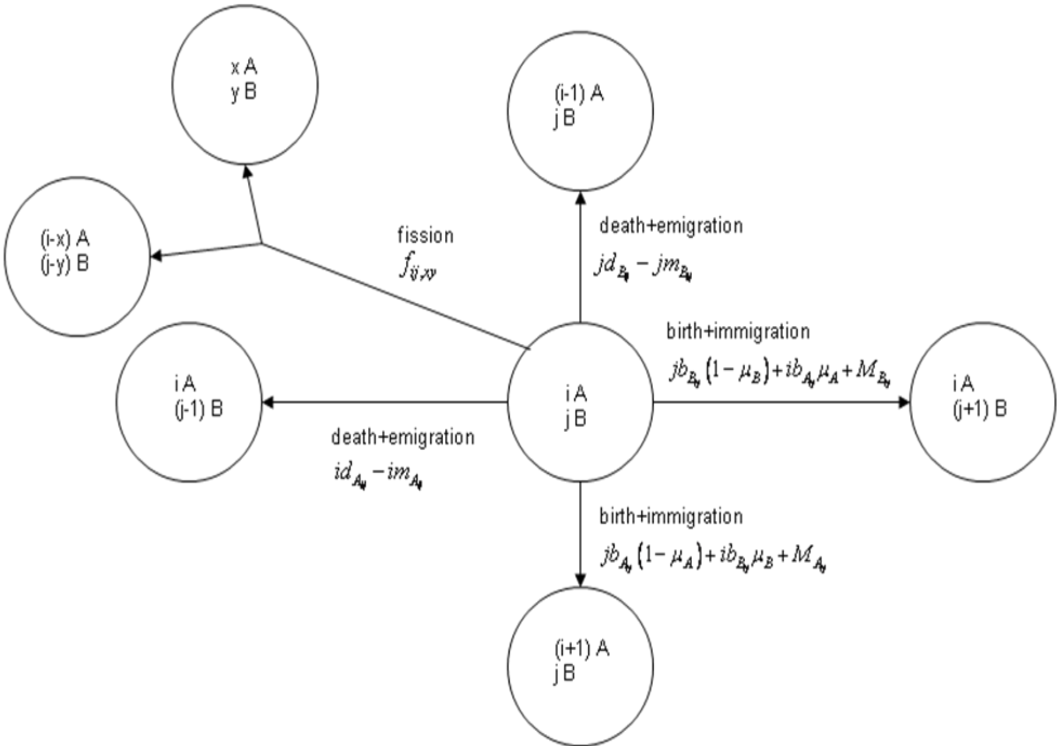


Figure 6. Representation of possible transitions from a group with *i* A-types and *j* B-types. Birth (*b*), death (*d*), emigration (*m*), and fission (*f*) rates depend on the individual and group type. Immigration (*M*) depends on the group distribution as well, since it depends on the number and type of individuals leaving other groups. There are many fission transitions (with just a general case shown), since there are many ways that a group can be divided (all combinations of *x* and *y*, in which $x \leq i$ and $y \leq j$). Fusion is not shown. Mutation (μ) is incorporated with births.

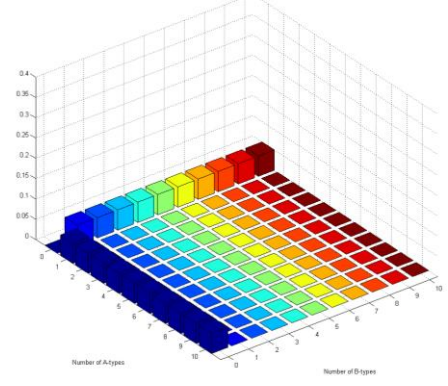
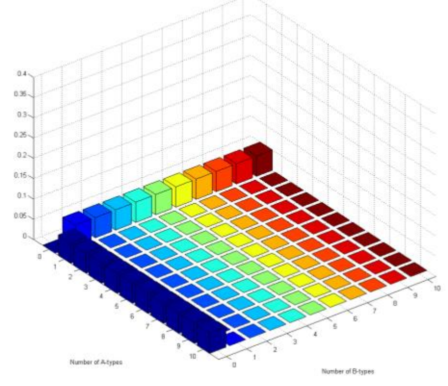
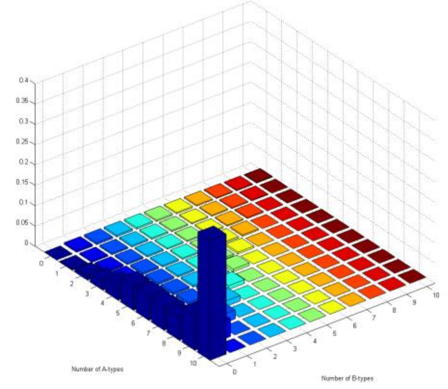
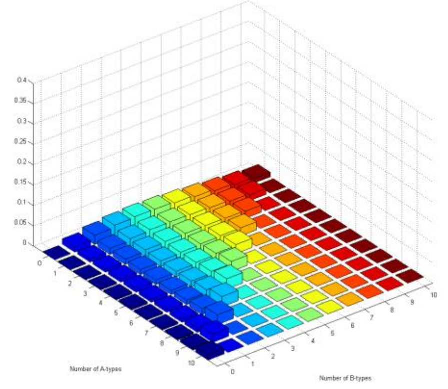
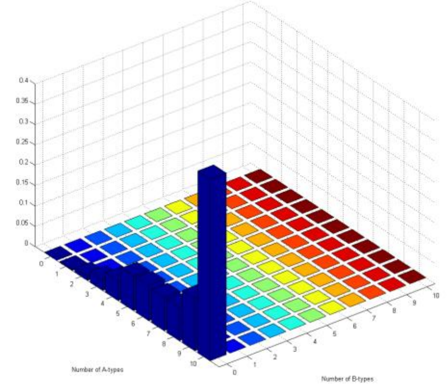
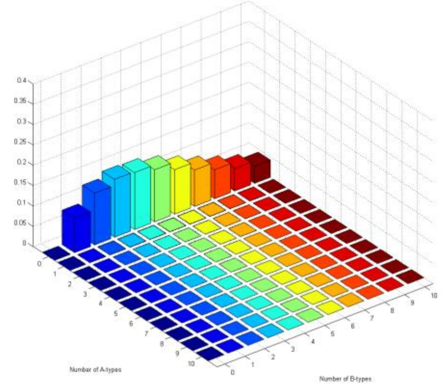
time	Low migration	High migration
0		
10		
100		

Figure 7. Group distribution changes through time for the case of classic group selection under low and high migration. Altruists are A-types, cheaters are B-types. Parameter values can be found in Appendix 1.

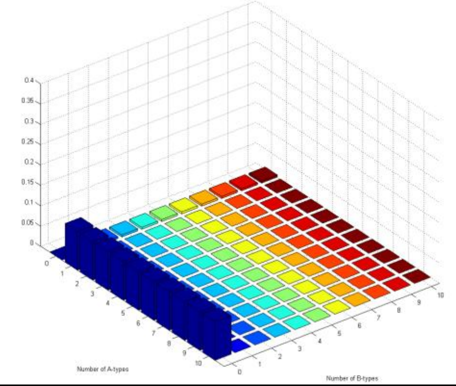
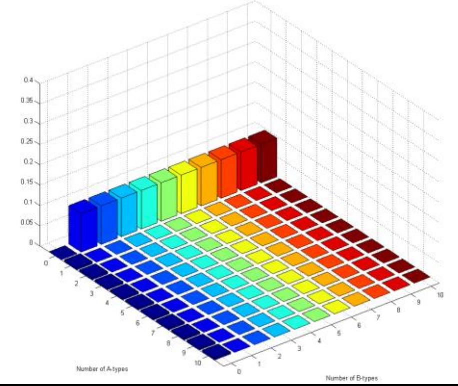
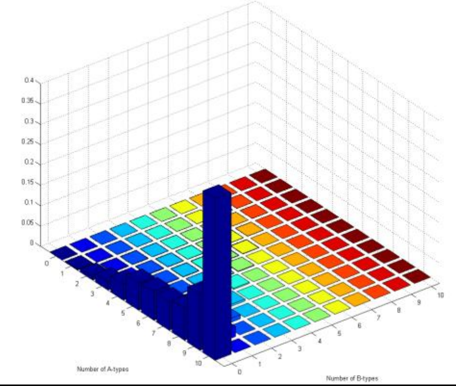
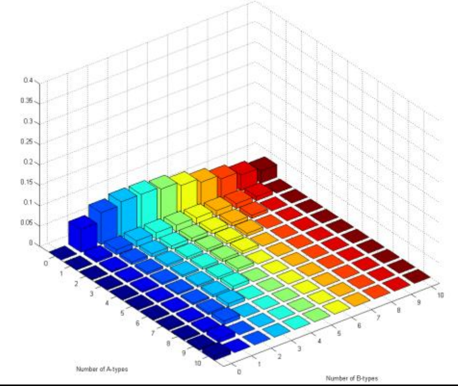
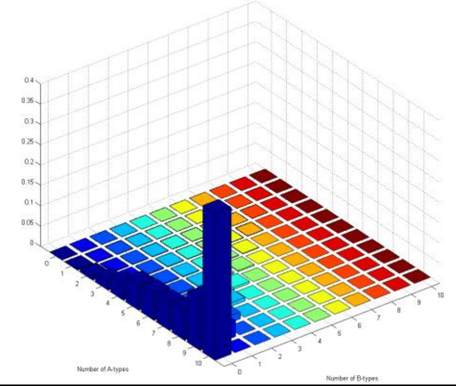
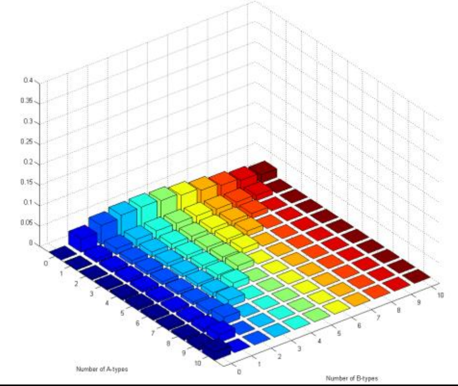
time	Initial A-dominant	Initial B-dominant
0		
10		
100		

Figure 8. Group distribution changes through time for the case of classic group selection with intermediate migration under different initial distributions. Altruists are A-types, cheaters are B-types. Parameter values can be found in Appendix 1.

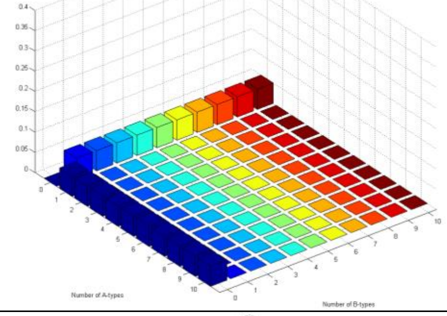
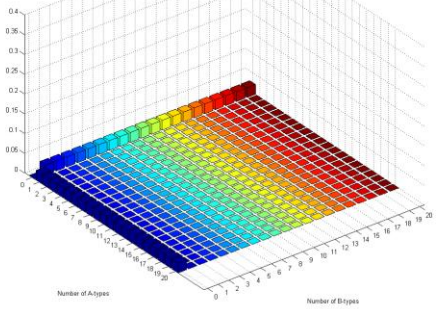
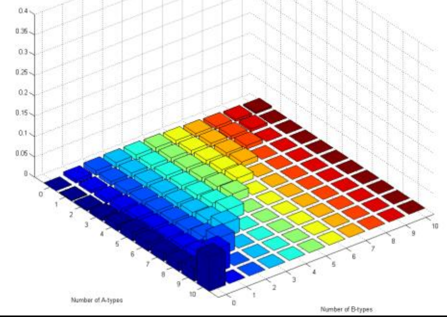
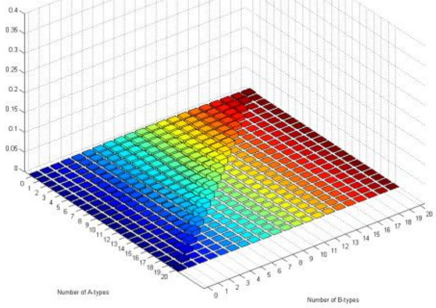
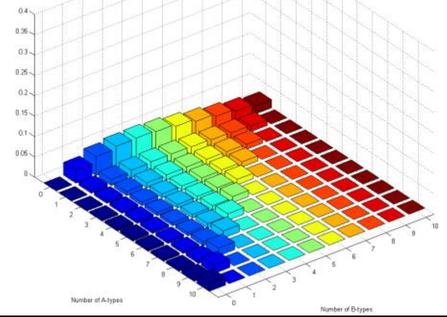
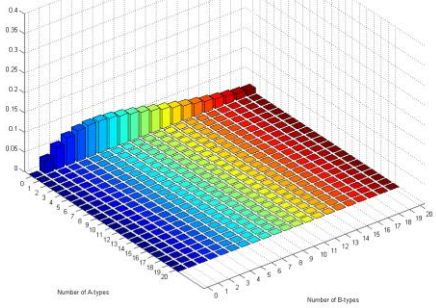
time	Maximum group size of 10	Maximum group size of 20
0		
10		
100		

Figure 9. Group distribution changes through time for the case of classic group selection with intermediate migration under different maximum population sizes. Altruists are A-types, cheaters are B-types. Parameter values can be found in Appendix 1.

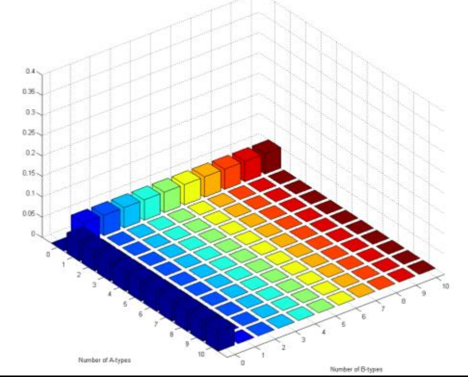
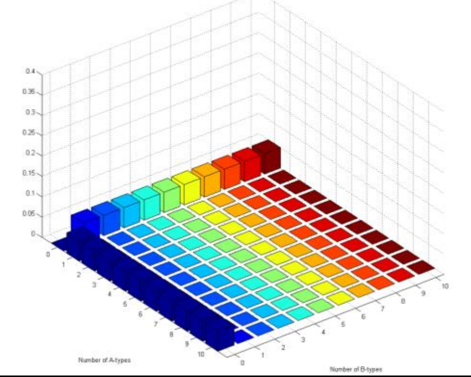
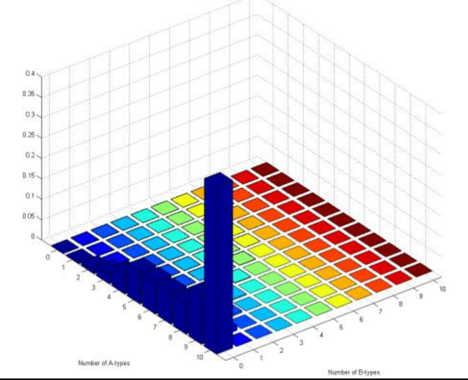
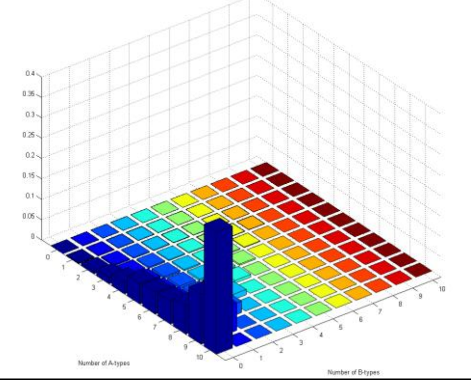
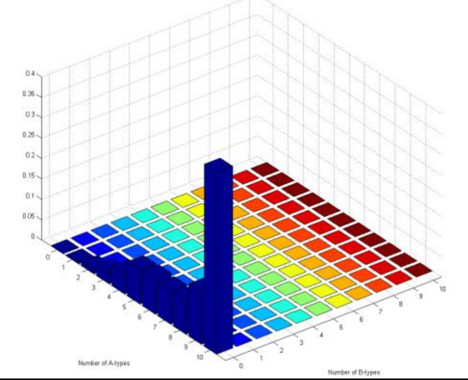
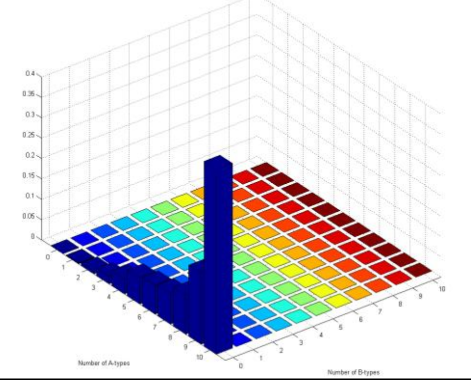
time	Low migration	High migration
0		
10		
100		

Figure 10. Group distribution changes through time for the case of asymmetric positive frequency-dependent selection under low and high migration. Groups of A-types have higher average fitness. Parameter values can be found in Appendix 1.

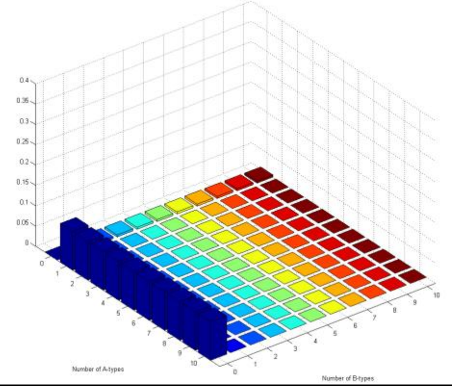
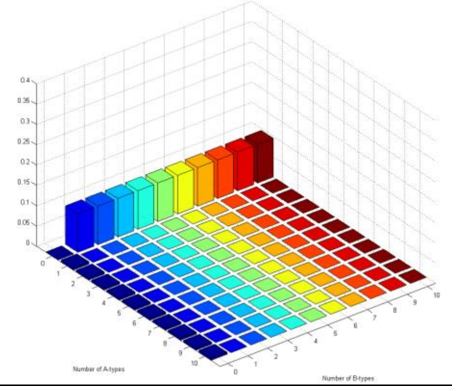
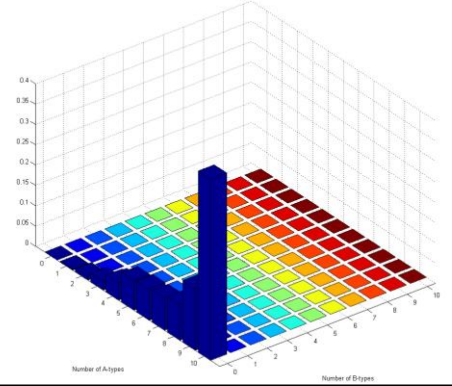
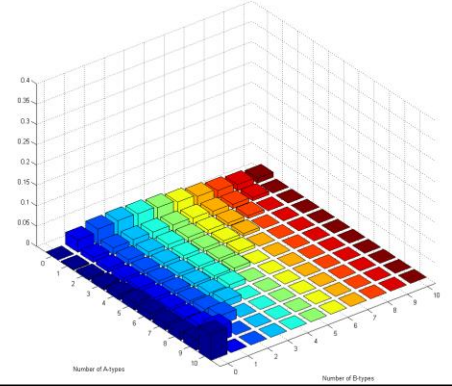
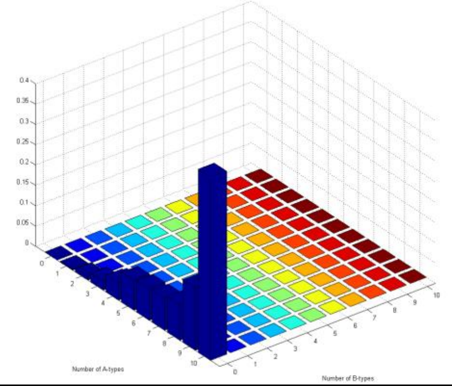
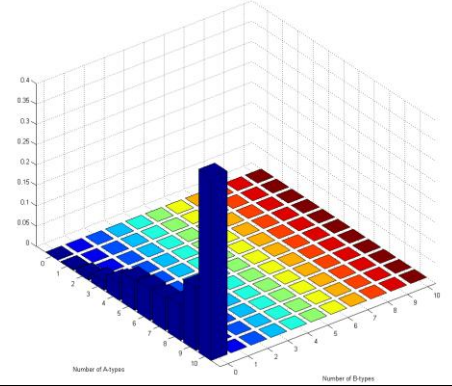
time	Initial A-dominant	Initial B-dominant
0		
10		
100		

Figure 11. Group distribution changes through time for the case of asymmetric positive frequency- dependent selection with intermediate migration under different initial distributions. Groups of A-types have higher average fitness. Parameter values can be found in Appendix 1.

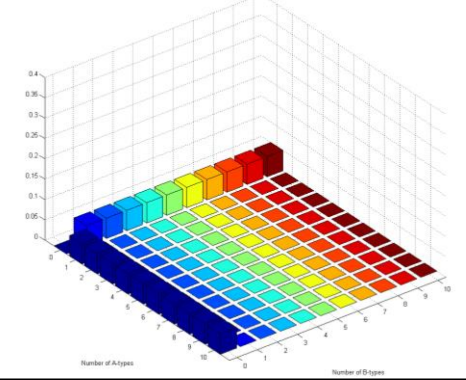
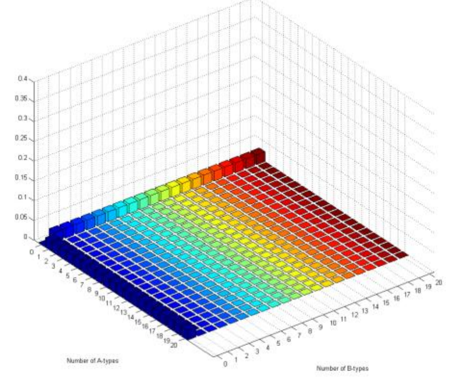
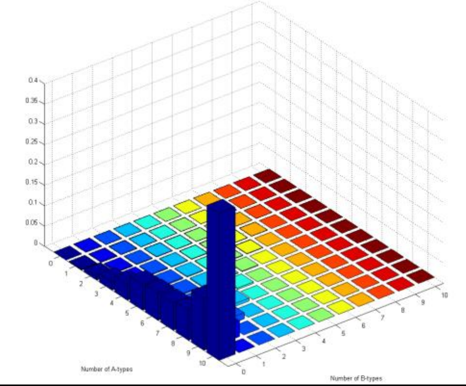
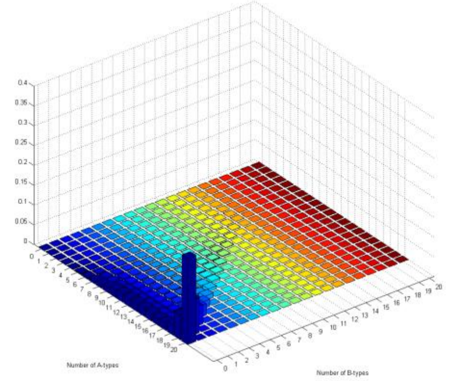
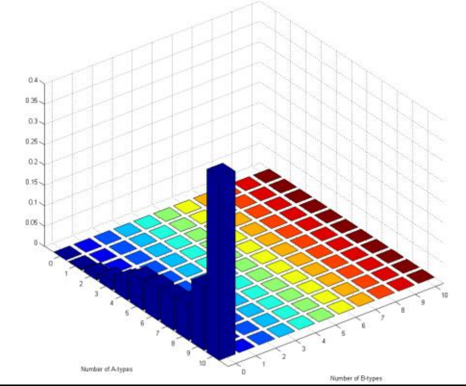
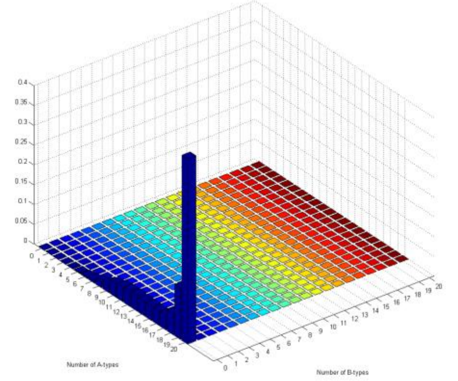
time	Maximum group size of 10	Maximum group size of 20
0		
10		
100		

Figure 12. Group distribution changes through time for the case of asymmetric positive frequency- dependent selection with intermediate migration under different maximum population sizes. Groups of A-types have higher average fitness. Parameter values can be found in Appendix 1.

5.2 Other spatial approaches

So far, we have only considered interacting discrete entities. However, it is also possible to model spatial structure using different approaches without detracting from the conceptual understanding of higher-level selection. Nowak and May (1992) modeled spatial games in which each individual occupies a discrete place on a spatial grid and interacts with individuals in neighboring spots. Mallet and Barton (1989) modeled positive frequency-dependence in a continuous population using diffusion equations. These approaches yield similar results. Though interestingly, the spread of group beneficial traits under strongly asymmetric positive frequency-dependence is treated as almost trivial. However, regardless of the specific spatial structure of the model, when lower-level fitness depends on some higher-level context and higher-level contexts have differential rates of persistence, whether their extent is measured as a range or as a number or groups, the concept of high-level selection is beneficial in understanding the resulting evolutionary dynamics.

5.3 Structuring on multiple levels and non-genetic heritability

So far, we have focused on simple haploid cases in order to have only one additional level of structure (*i.e.* group-level). However, structuring is often multilevel. With a standard simple diploid system, individual-level selection is reducible to gene-level selection. However, there are many cases in which even relatively simple diploid systems are not reducible. For example, overlapping generations, non-panmictic mating, epigenetics, and other additional considerations make it necessary to know how the genes are distributed into individuals in order to model the changes in gene frequencies, at least when the fitness of the gene depends on the individual type. Multiple levels, multiple genes, as well as non-genetic heritable traits and emergent properties, can all be incorporated within the multilevel framework developed here. For any level being modeled, one simply (at least on a theoretical level) requires that all possible types be fully described (*e.g.* by composition given by exact genotypes plus any cultural states) as well as the transitions between them. Since processes such as cultural evolution can be quite important in certain species, particularly humans (Richerson and Boyd 2004), the ability to incorporate them in addition to genetic processes into a single model is quite beneficial.

This even applies to higher levels. For example, communities may be defined by the species they contain (and one could also incorporate population sizes, genetics within populations, etc., though the number of community types quickly increases). Communities then transition to other community types through the loss or gain of species (and go extinct when they contain no species). Which species are introduced may depend on neighboring communities. Some communities may be stable and persist while others transition to other types or go extinct. New communities may be formed by colonization of empty areas from species (subsets) of neighboring communities. The dynamics are very much analogous to our discussion of groups and we can see that selection favors certain community types over others. The same logic could also be applied to ecosystems or perhaps even biospheres.

It should be noted, however, that higher-level interests are not always easily interpreted as averages of lower-level metrics of fitness, even in the case with groups. For example, consider a group that is more likely to persist if the growth rate is low, since this may avoid chaotic population dynamics that increase extinction rates. A low growth rate would increase the average fitness of the group over the long term by keeping it non-zero for at least some individuals, though the one-generation fitness may be lower for all individuals. Of course group fitness would still also depend on stability, so altruistic sacrifices for lower growth rate may not be favored, though individually favored competition in which interference reduces the population growth rate might.

Chapter VI: A universal view

6.1 Elimination acts on all levels

Since adaptation by natural selection is a process by which traits that are unable to persist are eliminated, the traits, at any level of organization, that remain can, in some sense, be considered adaptive. To understand the pattern of group types observed in nature, we must understand which group types are eliminated by group selection. Eq. 4 suggests that the fittest groups are those that are stable and have high average fitness. Thus, altruistic groups that are susceptible to invasion are eliminated. On the other hand, cheater groups that cause a "tragedy of the commons" resulting in extinction are also eliminated. In this case, if cheaters successfully invaded all altruist groups and subsequently went extinct, the species would be eliminated. Hence we expect to see neither susceptible altruists nor cheaters that cannot persist on their own in nature. Even in the absence of the spread of group-beneficial traits, higher-level selection is still occurring when there is elimination of group-detrimental traits.

Furthermore, constraints and tradeoffs that limit individually favored traits to those that result in stable higher levels can be considered adaptations. For example, a species incapable, due to physiological constraints, of producing individuals that are overly efficient predators, in that they would drive their prey species to extinction, would be able to persist, while a species without such constraints would not. That is to say, natural selection will favor species that have evolved ways to prevent the evolution of group detrimental traits. Hence, Wynne-Edwards' (1962) original argument is not without some merit.

The elimination of configurations that cannot persist has been termed "ecological exclusion" (Ginzburg and Colyvan 2004) or "ecological elimination" (Colyvan and Ginzburg 2010). As an example, Ginzburg et al. (2010) argue that the elimination of populations that exceed a certain population growth rate (the May threshold) experience chaotic population dynamics and consequently are more likely to go extinct, resulting in independence of the allometric relationship between maximum population growth rate and body size. Such elimination is an extreme form of higher selection in some ways analagous to selection against inviable individuals. While this is a very important process, it does not explain how natural selection can lead to solutions of higher group average fitness. However, higher-level selection is not limited to this extreme. A group type that is able to persist on its own can be eliminated due to competition with another group type capable of persisting on its own (*i.e.* both are viable, but one is evolutionarily inferior). We have seen this with the case of positive frequency-dependence. Furthermore, we have seen that certain traits can serve to reconcile individual and group interests. Hence, group selection can produce adaptive groups by selecting for effective configurations in which individual and group interests are resolved. Such groups have high fitness since they possess both favorable extinction/formation rates and high stability.

6.2 Conclusion

We began by asking whether group adaptations exist and whether anything is gained from a group-level approach. I believe that we can answer yes to both of these questions. We have demonstrated that, from a quantitative perspective, group selection is reducible to individual selection in only a few cases and that, in all other cases, a MLS2 approach is necessary (even when only interested in individual-level questions). Furthermore, a MLS2 approach is still correct, though not necessary, in reducible cases. We have also shown that, qualitatively, there exist situations in which a higher-level perspective is necessary to correctly predict evolutionary outcomes. Again, the higher-level approach remains correct, even when not necessary. Therefore, it may frequently be useful to record group types and their change over time when performing natural studies of evolution.

We have further argued that group selection is an elimination process that results in groups that are better able to persist, and thus one can reasonably interpret the outcomes of group selection as being adaptive. Higher-level adaptations are more than apparent; they are the actual result of a very real process.

References

Aviles, L. 1993. Interdemic selection and the sex ratio: a social spider perspective. American Naturalist, 142:320-345.

Boorman, S.A. and P.R. Levitt. 1980. The Genetics of Altruism. New York: Academic Press.

Boyd, R. and P.J. Richerson. 1990. Group selection among alternative evolutionarily stable strategies. Journal of Theoretical Biology, 145:331-342.

Choi, J. and S. Bowles. 2007. The coevolution of parochial altruism and war. Science, 318:636-640.

Colyvan, M. and L.R. Ginzburg. 2010. Analogical thinking in biology: looking beyond disciplinary boundaries. The Quarterly Review of Biology, 85(2):1-12.

Damuth, J. and I.L. Heisler. 1988. Alternative formulations of multilevel selection. Biology and Philosophy 3(4):407-430.

Darwin, C. 1871. The Descent of Man and Selection in Relation to Sex. New York: D. Appleton and Company.

Dawkins, R. 1976. The Selfish Gene. Oxford: Oxford University Press.

Dugatkin, L.A. and J.J. Godin. 1992. Reversal of femal mate choice by copying in the guppy (*Poecilia reticulate*). Proc. R. Soc. Lond. B. 249(1325):179-184.

Gilpin, M.E. 1975. Group Selection in Predator-Prey Communities. Princeton: Princeton University Press.

Ginzburg, L.R. and M. Colyvan. 2004. Ecological Orbits: How Planets Move and Populations Grow. Oxford: Oxford University Press.

Ginzburg, L.R., O. Burger, and J. Damuth. 2010. The May threshold and life-history allometry. Biology Letters, 6:850-853.

Glymour, B. 2008. Stable models and causal explanation in evolutionary biology. Philosophy of Science, 75:571-583.

Goodnight, C. J., J.M. Schwartz, and L. Stevens. 1992. Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. The American Naturalist, 140(5):743-761.

Hamilton, W.D. 1964. The genetical evolution of social behaviour I and II. Journal of Theoretical Biology, 7:1-16, 17-32

Hardin, G. 1968. Tragedy of the commons. Science, 162:1243-1248.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. 2007. Via freedom to coercion: the emergence of costly punishment, Science 316, 1905-1907.

Heisler, I.L. and J. Damuth. 1987. A method for analyzing selection in hierarchically structured populations. The American Naturalist, 130(4):582-602

Levin, B.R. and W.L. Kilmer. 1974. Interdemic selection and the evolution of altruism: a computer simulation study. Evolution, 28:527-545.

Levins, R. 1970. Extinction. In M. Gerstenhaber (ed.), Some Mathematical Problems in Biology. Providence: The American Mathematical Society.

Lewontin R.C. 1970. The units of selection. Annual Rev Ecol Syst 1:1–18.

Mallet, J. and N. Barton. 1989. Inference from clines stabilized by frequency-dependent selection. Genetics 122:967-976.

Maynard Smith, J. 1964. Group Selection. Nature, 201:1145-1147.

Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge: Cambridge University Press.

Maynard Smith, J. 1999. Conflict and cooperation in human societies. In L. Keller (ed.), Levels of Selection in Evolution. Princeton: Princeton University Press.

McCauley, D.E. 1994. Intrademic group selection imposed by a parasitoid-host interaction. The American Naturalist, 144(1):1-13.

Nash, J. 1950. Equilibrium points in n-person games. Proceedings of the National Academy of Sciences 36(1):48-49.

Nowak M.A., R.M. May. 1992. Evolutionary games and spatial chaos. Nature, 359: 826-829.

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

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

Pollock, G.B., A. Cabrales, and S.W. Rissing. 2004. On suicidal punishment among *Acromyrmex versicolor* co-foundresses: the disadvantage in personal advantage. Evolutionary Ecology Research 6:891-917.

Price, G.R. 1972. Extension of covariance selection mathematics. Annals of Human Genetics, 35(4):485–490.

Richerson, P.J. and R.Boyd. 2004. Not by genes alone: how culture transformed human evolution. University of Chicago Press.

Rissing, S.W., G.B. Pollock, M.R. Higgins, R.H. Hagen, and D.R. Smith. 1989. Foraging specialization without relatedness or dominance among co-founding ant queens. Nature, 338:420-422.

Roughgarden, J. 1998. Primer of Ecological Theory. Upper Saddle River, New Jersey: Prentice Hall.

Sober, E. and D. S. Wilson. 1998. Unto Others: The Evolution and Psychology of Unselfish Behavior. Cambridge, MA.: Harvard University Press.

Taylor, C. and M. Nowak. 2007. Transforming the dilemma. Evolution 61(10):2284-2292.

Trivers, R.L. 1971. The evolution of reciprocal altruism. The Quarterly Review of Biology, 46: 35-57.

Weatherhead, P.J. and R.J. Robertson. 1979. Offspring quality and the polygyny threshold: 'The sexy son hypothesis'. American Naturalist 113 (2): 201–8.

Williams, G.C. 1966. Adaptation and Natural Selection. Princeton: Princeton University Press.

Wilson, D. S. 1975. A theory of group selection. Proceedings from the National Academy of Sciences, 72: 143-146.

Wynne-Edwards, V. C. 1962. Animal Dispersion in Relation to Social Behaviour. Edinburgh: Oliver and Boyd Publishing.

Wright, S. 1980. Genic and organismic selection. Evolution, 34:825-843.

Appendix 1. Parameter values used for Figures 7-12.

Fixed parameters:

birth rate	1
fission rate	0.693
mutation rate	0.0001
slope A	2
intercept A	0.5
intercept B	1

Altered parameters:

Selection type	**Figure**	**Condition**	**slope B**	**migration rate**	**% A groups**	**% B groups**	**max group size**
Classic group selection	7	Low migration	2	0.105	50	50	10
		High migration	2	0.357	50	50	10
	8	Initial A	2	0.223	95	5	10
		Initial B	2	0.223	5	95	10
	9	Max size 10	2	0.223	50	50	10
		Max size 20	2	0.223	50	50	20
Asymmetric positive frequency-dependent selection	10	Low migration	1	0.105	50	50	10
		High migration	1	0.357	50	50	10
	11	Initial A	1	0.223	95	5	10
		Initial B	1	0.223	5	95	10
	12	Max size 10	1	0.223	50	50	10
		Max size 20	1	0.223	50	50	20

Appendix 2. Matlab code for running the model for Figures 7-12.

Model.m:

```
clear all
tic
dt=.01;%small dt to approach continuous
t=100;%0,10,100
n=10;%max group size
G0=zeros(n+1);
G0(1,2:n+1)=0.01*50/n;
G0(2:n+1,1)=0.01*50/n;%50:50, 95:5, or 5:95
V0=(reshape(transpose(G0),1,(n+1)^2));
G=G0;
V=V0;
b=1;
mig=-log(1-.1);%low=.1, med=.2, high=.3,
fiss=-log(1-0.5);
mut=0.0001;
slopeA=2;
interceptA=.5;
slopeB=2;%2 for classic 1 for positive frequency-dependence
interceptB=1;

for j=0:n
    for k=0:n
        if j+1<=n

B((j)*(n+1)+(k)+1,(j+1)*(n+1)+(k)+1)=b*j*(j+k+1<=n);%birth A with
ceiling density dependence
        B((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=-b*j*(j+k+1<=n);
        else B((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end

        if k+1<=n

B((j)*(n+1)+(k)+1,(j)*(n+1)+(k+1)+1)=b*k*(j+k+1<=n);%birth B with
ceiling density dependence

B((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=B((j)*(n+1)+(k)+1,(j)*(n+1)+(k
)+1)-b*k*(j+k+1<=n);
        end

        if j-1>=0
        D((j)*(n+1)+(k)+1,(j-1)*(n+1)+(k)+1)=(b-
log(interceptA+slopeA*((j-1)/max((j+k-1),1))))*(j);%death A
        D((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=-(b-
log(interceptA+slopeA*((j-1)/max((j+k-1),1))))*(j);
        else D((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end
```

```
        if k-1>=0
        D((j)*(n+1)+(k)+1,(j)*(n+1)+(k-1)+1)=(b-
log(interceptB+slopeB*((j)/max((j+k-1),1))))*(k);%death B

D((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=D((j)*(n+1)+(k)+1,(j)*(n+1)+(k
)+1)-(b-log(interceptB+slopeB*((j)/max((j+k-1),1))))*(k);
        end

        if j-1>=0
        E((j)*(n+1)+(k)+1,(j-1)*(n+1)+(k)+1)=mig*(j);%A emigrate
        E((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=-mig*(j);
        else E((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end

        if k-1>=0
        E((j)*(n+1)+(k)+1,(j)*(n+1)+(k-1)+1)=mig*(k);%B emigrate

E((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=E((j)*(n+1)+(k)+1,(j)*(n+1)+(k
)+1)-mig*(k);
        end

        for x=0:j
        for y=0:k
        F((j)*(n+1)+(k)+1,(x)*(n+1)+(y)+1)=0;
        F((j)*(n+1)+(k)+1,(j-x)*(n+1)+(k-
y)+1)=2*fiss*binopdf(x,j,.5)*binopdf(y,k,.5)*(j+k==n);%group
fissions, only happens when reaches carrying capacity, fiss rate
includes 0,n

F((j)*(n+1)+(k)+1,(x)*(n+1)+(y)+1)=F((j)*(n+1)+(k)+1,(x)*(n+1)+(y
)+1)+2*fiss*binopdf(x,j,.5)*binopdf(y,k,.5)*(j+k==n);
        end
        end

F((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=F((j)*(n+1)+(k)+1,(j)*(n+1)+(k
)+1)-fiss*(j+k==n);

        if j+1<=n

IA((j)*(n+1)+(k)+1,(j+1)*(n+1)+(k)+1)=mig*(j+k+1<=n)*(j+k>0);%rec
ieve A immigrant (need to multiply by A density)
        IA((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=-
mig*(j+k+1<=n)*(j+k>0);
        else IA((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end

        if k+1<=n
```

```
IB((j)*(n+1)+(k)+1,(j)*(n+1)+(k+1)+1)=mig*(j+k+1<=n)*(j+k>0);%rec
ieve B immigrant (need to multiply by B density)

IB((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=IB((j)*(n+1)+(k)+1,(j)*(n+1)+
(k)+1)-mig*(j+k+1<=n)*(j+k>0);
        else IB((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end

        Grec(j+1,k+1)=1*(j+k+1<=n)*(j+k>0);%to determine the
number of groups that recieve the migrants, each group here gets
equal weight

        if j+1<=n
        M((j)*(n+1)+(k)+1,(j+1)*(n+1)+(k)+1)=-
mut*b*j*(j+k+1<=n);%what would have been A no longer goes there
        if k+1<=n

M((j)*(n+1)+(k)+1,(j)*(n+1)+(k+1)+1)=mut*b*j*(j+k+1<=n);%goes to
B instead
        end
        else M((j)*(n+1)+(k)+1,(j)*(n+1)+(k)+1)=0;
        end

        if k+1<=n
        M((j)*(n+1)+(k)+1,(j)*(n+1)+(k+1)+1)=-
mut*b*k*(j+k+1<=n);%what would have been B no longer goes there
        if j+1<=n

M((j)*(n+1)+(k)+1,(j+1)*(n+1)+(k)+1)=M((j)*(n+1)+(k)+1,(j)*(n+1)+
(k)+1)+mut*b*k*(j+k+1<=n);%goes to A instead
        end
        end
    end
end

C=B+M+D+E+F;
for i=2:t/dt+1
    V=Vf(C,G,Grec,n,IA,IB,V,dt);
    G=transpose(reshape(V,n+1,n+1));
end
Gt=transpose(reshape(V,n+1,n+1));
Gtdist=Gt;
Gtdist(1,1)=0;
Gtdist=Gtdist/sum(sum(Gtdist))
bar3(Gtdist),xlabel('Number of B-types'),ylabel('Number of A-
types'),
set(gca,'XTickLabel',0:n),set(gca,'YTickLabel',0:n),set(gca,'XTic
k',1:n+1),set(gca,'YTick',1:n+1)
toc
```

Vf.m:

```
function [Vf]= Vf(C,G,Grec,n,IA,IB,V,dt)
    numA=cumsum([0 ones(1,n)])*G*(ones(n+1,1));
    numB=ones(1,n+1)*G*cumsum([0;ones(n,1)]);
    grec=ones(1,n+1)*(G.*Grec)*ones(n+1,1);%since only a portion
of groups receive migrants
    r=C+numA/grec*IA+numB/grec*IB;
Vf=V*expm(r*dt);
end
```