UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
School of Electronics and Computer Science

Social Niche Construction: Evolutionary Explanations for Cooperative Group Formation

by
Simon T. Powers

Thesis for the degree of Doctor of Philosophy

October 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

SOCIAL NICHE CONSTRUCTION: EVOLUTIONARY EXPLANATIONS FOR COOPERATIVE GROUP FORMATION

by Simon T. Powers

Cooperative behaviours can be defined as those that benefit others at an apparent cost to self. How these kinds of behaviours can evolve has been a topic of great interest in evolutionary biology, for at first sight we would not expect one organism to evolve to help another. Explanations for cooperation rely on the presence of a population structure that clusters cooperators together, such that they enjoy the benefits of each others' actions. But, the question that has been left largely unaddressed is, how does this structure itself evolve? If we want to really explain why organisms cooperate, then we need to explain not just their adaptation to their social environment, but why they live in that environment.

It is well-known that individual genetic traits can affect population structure; an example is extracellular matrix production by bacteria in a biofilm. Yet, the concurrent evolution of such traits with social behaviour is very rarely considered. We show here that social behaviour can exert indirect selection pressure on population structure-modifying traits, causing individuals to adaptively modify their population structure to support greater cooperation. Moreover, we argue that any component of selection on structure-modifying traits that is due to social behaviour must be in the direction of increased cooperation; that component of selection cannot be in favour of the conditions for greater selfishness. We then examine the conditions under which this component of selection on population structure exists. Thus, we argue that not only can population structure drive the evolution of cooperation, as in classical models, but that the benefits of greater cooperation can in turn drive the evolution of population structure – a positive feedback process that we call *social niche construction*.

We argue that this process is necessary in providing an adaptive explanation for some of the major transitions in evolution (such as from single- to multi- celled organisms, and from solitary insects to eusocial colonies). Any satisfactory account of these transitions must explain how the individuals came to live in a population structure that supported high degrees of cooperation, as well as showing that cooperation is individually advantageous given that structure.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, *Simon T. Powers*, declare that the thesis entitled *Social Niche Construction: Evolutionary Explanations for Cooperative Group Formation* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: Powers et al. (2007, 2008, 2009); Powers and Watson (2008, 2009).

Signed:..
Date: 27/09/2010...

Acknowledgements

I must first thank my supervisor, Richard Watson, for giving me complete freedom to pursue this research direction. His advice and enthusiasm for my research has been invaluable.

Alex Penn started me along the path of this research direction, by bringing the possibility of evolvable traits affecting the strength of group selection to my attention. My research has benefited from many discussions with her.

I have enjoyed many interesting discussions over lunch and coffee with members of the Science & Engineering of Natural Systems (SENSe) Group here at Southampton (Alex, Chris, Jason, Johannes, Nic, Richard, Seth, Rob, Rob and Tom). Thanks in particular to Seth Bullock for very helpful advice, and Jason Noble for constructive feedback during my progress examinations. The SENSe postgraduate lab has been a great place to work, with great company. A big thanks to all of my friends there for the support (both moral and practical) that they have given me.

Finally, I must thank my parents for all of the love and support they have given me in this, and everything else that I have done.

Chapter 1

Introduction

The question of how cooperative social behaviours between organisms can evolve has troubled evolutionary biologists since Darwin. If the maxim of natural selection is "survival of the fittest", then why should one organism perform an act that helps, and therefore increases the fitness of, another?[1] This accords with the popular view of "nature red in tooth and claw", with raw competition as the driving force in evolution. Despite this, however, examples of organisms cooperating and helping each other can readily be found in nature across all taxa. For example, microbes share public goods that they produce (Buckling et al., 2007; Gore et al., 2009), such as enzymes, vampire bats share blood with those in their nest that failed to feed themselves (Wilkinson, 1984), and many mammals look after each others young. Moreover, while these macroscopic examples of cooperation can easily be observed, cooperation has also been the key driver for the creation of all new levels of organisation in the biological hierarchy, from protocells through to modern day multicellular organisms and social insect colonies (Buss, 1988; Maynard Smith and Szathmáry, 1995; Michod, 1999). Cooperation has, therefore, played a fundamental role in the creation of all life on earth, despite its *prima facie* incompatibility with Darwinian logic, and is still at work inside of all modern day organisms. As an example, animals have a separation between germ-line and somatic cells: only the germ-line cells are passed on to the offspring multicellular organism; the somatic cells, such as skin, cooperate by forgoing their own direct reproduction to contribute to the success of the group of cells (i.e., the multicellular organism) as a whole.

The key difficulty in explaining all of the examples of cooperation above is the problem of selfish cheats. These are individuals that do not themselves cooperate, but that do still benefit from the help of those around them. Thus if cooperating carries a fitness cost, such as time or energy that could have been spent directly on survival or reproduction, then cheats should be fitter than cooperators, leading natural selection to ultimately drive the cooperative type extinct. Referring back to the above examples, instances

---

[1] We give more precise definitions of cooperative social behaviours later in this chapter.

of cheating would be microbes that do not produce enzymes themselves but just use those supplied by others, vampire bats that do not share blood but still receive it from others, or mammals that do not care for others' young, but allow others to care for their own. In the case of cooperation within multicellular organisms, examples of cheating are particularly vivid: they are cancerous somatic cells that do not curtail their own replication rate for the good of the other cells in the organism. All of these kinds of cheating behaviour could readily arise in nature as new mutations. In particular, it is often suggested that it is easier to lose a cooperative behaviour than to gain one, since cheating behaviour may arise simply by the loss of a gene that codes for the cooperative act. The possibility of cheating creates a social dilemma (Dawes, 1980) in which *absolute* fitness would be higher if all individuals cooperated, but cheating confers a *relative* advantage, even though it leads to an overall decline in mean fitness.

In this chapter, we first consider the range of cooperative behaviours that occur in nature, and then look at what is already known about the conditions under which they can evolve. Essentially, these conditions boil down to aspects of the organisms' population structure, such as the mating system (Wade and Breden, 1981), the distance offspring disperse from their parents (West et al., 2002), and the size of social groups (Wilson, 1983b). What is very rarely considered though, is how this population structure itself evolves, for it is usually just assumed to be a part of the organisms' physical environment. Thus, cooperation is then simply the adaptation of an organism's social behaviour to its current environment. But, population structure can evolve over time, and so the selection pressures on social behaviour can in fact change. Moreover, we contend that this change can happen in an adaptive manner, for individual genetic traits can affect population structure, and can themselves be subject to selection alongside social behaviour. Thus, the joint equilibrium of structure and behaviour that a given population has reached needs an explanation over and above classical explanations for the level of cooperation in a fixed structure, once we treat the structure as a factor endogenous to evolution. That is, why is the equilibrium at a structure that supports x level of cooperation rather than y? Our thesis seeks to provide an adaptive explanation for why this equilibrium is where it is on the axes of both behaviour and structure.

We will advance the thesis that any selection pressure on population structure, derived through its effect on social behaviour, must be towards the creation of selective environments favouring cooperation; the conditions for greater selfish behaviour cannot be selectively favoured *per se*. While most work on social evolution shows that population structure can drive the evolution of cooperation, we argue that cooperation will also tend to drive the evolution of population structure. In particular, our work can explain the *origin* of groups that support cooperation, for while most models assume a group structure, we show how the concurrent evolution of population structure and social behaviour leads to the creation of social groups. This speaks directly to the group selection controversy, that is, on whether groups can be subject to natural selection in the same way that individual organisms are (Darwin, 1871; Haldane, 1932; Wright, 1945; Wynne-Edwards, 1962; Wilson, 1980; Sober and Wilson, 1998). There exists a strong individualist tendency in the literature, dating back to the 1960s, that sidelines the evolutionary role of groups (Williams, 1966; Dawkins, 1976, 1994; West et al., 2008). Yet, we show here how individuals create the very conditions in which group selection becomes an effective force. Our thesis suggests that when the groups are being adaptively created by the organisms themselves, then they should not simply be viewed as part of the selective background, for their very creation can change both the sign and magnitude of selection pressure on social behaviour.

Our thesis is also particularly relevant to the major transitions in evolution (Maynard Smith and Szathmáry, 1995), such as from single- to multi- celled organisms, and from solitary organisms to societies. These transitions involved the creation of groups whose members exhibit a high degree of cooperation, so much so that the group becomes so well integrated that we call it an evolutionary individual in its own right (Maynard Smith and Szathmáry, 1995; Michod, 1999). That is, the equilibrium of behaviour and structure moves towards the extreme end of both axes during a transition. For example, the population of cells in a multicellular organism is so well structured that the individual cells cannot now live a solitary existence, and the level of cooperation is consequently exceptionally high. Our thesis can provide a high level explanation for how the equilibrium can move from being one of loosely associated groups with little cooperation, to one where the groups are so rigid that their members cannot reproduce outside of them, and cooperation is very high. Moreover, we explain how this movement of the equilibrium is the result of natural selection, and hence is individually adaptive. In light of this, we can provide the basis of a systematic selective account that is general to several of the transitions.

In a sense, our thesis can explain the evolution of cooperation in a more fundamental way than most previous works. This is because a behaviour is only "cooperative" relative to something. In most models, that something is the null hypothesis that the behaviour would not evolve in a freely-mixed population. But if it is a fact of life that the organisms under consideration do not live in a freely-mixed population, then this is something of a meaningless comparison. The evolution of cooperation shown in such models is then simply the cooperative allele reaching its equilibrium frequency under the given population structure. Thus, the organisms are not really becoming more cooperative, for social behaviour is just reaching its equilibrium. But if the population structure can itself evolve, then we can have a meaningful measure of cooperation increase. This measure is how far the equilibrium of behaviour has moved as a result of movement in population structure. We can therefore say that the organisms are more cooperative relative to the original population structure that they actually used to live in, rather than to the hypothetical freely-mixed case. In that sense, we show a meaningful evolutionary increase in cooperation.

1.1 Defining cooperation

The term “cooperation” has a variety of different meanings in the evolutionary biology literature. All definitions generally agree that cooperative behaviours should have a positive effect on the fitness of recipients. A fundamental issue is then whether or not this positive effect on the recipient comes at some cost to the actor. In this dissertation, we mainly consider cooperative behaviours that carry a *relative* fitness cost to the actor, that is, they increase the fitness of the recipient(s) by more than they increase the fitness of the actor. As discussed below, these traits include “public goods production”, whereby an individual provides a benefit to its whole group (including itself), such as an extracellular enzyme or an alarm call that a predator is approaching. Such public goods, however, are typically costly for the individual to produce, and hence put the actor at a fitness disadvantage compared to non-producers in the same group. These kinds of traits have been termed “weakly altruistic” by Wilson (1979, 1980, 1990), since the fitness cost incurred by the actor is *relative* to other group members, rather than absolute (i.e., the trait may increase the absolute number of offspring of the actor, whilst increasing the number of offspring of other group members by even more)[2]. The models presented in chapters 3, 4 and 6 are all concerned with these kinds of traits that are beneficial to the whole group, but at a relative cost to the actor. In Chapter 5 we consider other forms of social behaviour, which have also been called cooperative in the literature. In these cases, cooperation involves coordination of behaviour in order to achieve a task that cannot be completed alone. An example of such behaviour is “cooperative” hunting in packs of lions (Packer and Ruttan, 1988). By contrast, the “weakly altruistic” traits that we consider represent activities which can be done alone, but that benefit other group members.

1.1.1 Examples of cooperative behaviour: public goods and the Tragedy of the Commons

What kinds of behaviour are cooperative, then, in the sense that they carry a fitness cost for the actor? Perhaps the pinnacle examples of behaviours that increase the immediate fitness of recipients by more than the actor are those altruistic traits that involve either suicide or sterility on the part of the bearer. Indeed, it was the sterility of workers in social insect colonies that first troubled Darwin (1859). A similar sterility is also seen in most of the cells in multicellular organisms that have a germ-soma separation (Buss, 1988; Maynard Smith and Szathmáry, 1995), for the somatic cells (skin,liver,etc.) are not passed on to the next generation of multicellular organisms. The suicidal sting of

---

[2] Strongly altruistic traits are those that decrease the *absolute* fitness of the actor, whilst increasing that of other group members. However, the distinction between strong and weak altruism is not fundamental if groups stay together for multiple organismic generations, since in that case both types of trait exhibit qualitatively similar selective dynamics (Fletcher and Zwick, 2004; Fletcher et al., 2006; Fletcher and Zwick, 2007). We discuss this point in detail in Chapter 2.

a honey bee provides another example of altruism, and such extreme cooperative acts as sterility and suicide has prompted some to view social insect colonies as "superorganisms" (Wheeler, 1911; Wilson and Sober, 1989; Hölldobler and Wilson, 2009). All of these behaviours can benefit other individuals: by foregoing longterm reproductive prospects, somatic cells are able to specialise in maintaining the survival of the multicellular organism (Michod, 1997); by attacking an intruder to the nest, a suicidal bee is protecting his nest mates. It is also clear how all of these behaviours could be exploited by selfish cheats: cheats would be otherwise identical organisms that did not perform the suicidal act or that did not forego their own reproduction, but that still received the benefits from others doing so.

In this dissertation, we focus on less extreme behaviours than sterility and suicide, for these are at the end of a fitness cost-to-self versus fitness benefit-to-others scale of cooperativeness. One example of the kind of cooperation that we consider is an organism's resource exploitation strategy. In particular, we consider cases where there exists a trade-off between growth rate and yield, i.e., the faster an organism consumes a resource, the lower the biomass produced per unit consumed. Such a scenario is more generally a case of the "Tragedy of the Commons" (Hardin, 1968; Rankin et al., 2007), or the problem of prudent predation (Slobodkin, 1964; Gilpin, 1975). This is because it is of immediate benefit to the individual to grow as fast as possible, and hence to consume the resource as fast as possible. Doing so, however, decreases yield such that the organism, and all others sharing the same resource, are worse off in the long term. Decreasing consumption rate would increase the total amount of biomass that the resource could support, but would require an organism to grow at a slower rate, which would put it at a fitness disadvantage compared to unrestrained resource users. In other words, although restrained resource usage is the optimal outcome for both the group, and the individual in the long term, it can be hard to see how unrestrained users can be kept in check, since they enjoy a local competitive advantage. It then follows that unrestrained users can be seen as "selfish", because they maximise their own short-term benefit to the detriment of the long-term well-being of the group, while restrained users can be seen as "cooperative", because they forgo this immediate short-term benefit in order to provide more resource for the group as a whole.

Under the classical model of natural selection in a freely-mixed population (Fisher, 1930; Haldane, 1932) selfish individuals must win out, for natural selection is sensitive only to immediate benefits - it does not have "foresight" and so cannot be sensitive to the long-term benefits of efficiency (Williams, 1966; Dawkins, 1976). Examples of where such resource utilisation trade-offs occur in nature include the virulence with which parasites attack their hosts (Bull et al., 1991; Williams and Nesse, 1991; Frank, 1996; Kerr et al., 2006), competition for water amongst plants (Zea-Cabrera et al., 2006), and the opportunity for production of adenosine triphosphate in heterotrophic organisms (Pfeiffer et al., 2001; Kreft, 2004; MacLean, 2008) through either respiration alone (slow but efficient), or through fermentation in addition to respiration (fast but inefficient). In all of these cases, short-term individual interests clash with the long-term interests of the group.

Another example of cooperative behaviours considered here concerns the production of public goods, that is, beneficial products that are manufactured by one individual but that are shared with all those nearby. For example, bacteria need iron to be able to grow, and in harsh environments they can sequester iron by producing scavenging molecules known as siderophores, that bind to iron and assist with its uptake (West and Buckling, 2003; Varma and Chincholkar, 2007). These siderophores are costly to produce, in terms of time and energy, yet are external substances that can be used by not just the producers, but by all bacteria that are nearby. As a result, a selection pressure for cheating behaviour exists, whereby individuals simply make use of the siderophores provided by their neighbours, without themselves incurring the metabolic costs (Buckling et al., 2007). Such cheats are therefore fitter than producers in the same locality, even though the whole group will do better the more producers are present (Griffin et al., 2004). Cooperation (producing) might therefore again be expected to break down in face of the immediate individual benefits of selfish behaviours.

It should be noted that public goods do not have to be physical substances, as in this example; they could be information. For example, in schools of guppy fish, one guppy will often swim up to and inspect a potential predator, and in so doing warn the whole school of danger (Dugatkin, 1990); similar behaviour can also be found in flocks of birds (Charnov and Krebs, 1975). There is, however, a crucial difference between sharing information and sharing physical substances such as food. This is that when an individual shares a physical substance, it loses precisely the amount that it shares. By contrast when an individual shares information, such as the location of a predator, that information does not immediately loose value to the sharer (Lachmann et al., 2000). That is, the *per capita* value of information does not decrease as the number of individuals it is shared with increases, in the way that the value of a physical substance does. Thus, we might expect the selective dynamics of information sharing within groups to be different from the dynamics of sharing physical substances. Indeed, Lachmann et al. (2000) suggest that the evolution of mechanisms for sharing information could drive the evolution of differentiation and specialisation during the creation of new levels of biological organisation, such as multicellular organisms and eusocial insect colonies.

A resource consumption strategy might at first sight seem to be a quite different problem to that of public goods production. However, the scenarios are in fact equivalent from the point of view of social evolution theory. Specifically, both types involve public goods: the first type involves their exploitation, the second their production. Likewise, both scenarios can lead to a Tragedy of the Commons: in the first case, the resource becomes overexploited and its benefit declines, in the second case, the manufactured public good is the resource, and its benefit too declines as cooperative producers are replaced by selfish non-producers. Also common is the fact a cooperative act by one individual is assumed to benefit the whole group, for the group is nearly always defined in social evolution models as the set of individuals that share the public good. Finally, in both cases, the actor also receives some benefit from its action; these are what Pepper (2000) has called "whole-group traits".

1.1.2 An evolutionary game theory perspective

The types of cooperation discussed above create a social dilemma that can be conceptualised by a payoff matrix. Such a matrix considers the interaction between a pair of individuals, and shows the fitness modifier that an individual will receive based on its own social strategy (cooperative or selfish) and that of its partner. The public goods scenarios that we have described above can be modelled using the payoff matrix shown in Table 1.1. In this matrix, b represents the benefit of cooperating, and c the cost. The equilibrium frequency of behaviours that will be reached depends upon the relationship between b and c. If $0.5 < b/c < 1$ then the game is a Prisoner's Dilemma (Doebeli and Hauert, 2005). This means that in a freely-mixed population, where individuals interact randomly, then the equilibrium is every individual adopting selfish behaviour. On the other hand, if $b/c > 1$ then the social interaction is a Snowdrift game (Sugden, 1986; Doebeli and Hauert, 2005). In that case, some stable ratio of cooperative and selfish individuals is reached at equilibrium in a freely-mixed population. The public goods scenarios that we have discussed can take either of these forms. Which they take depends on whether or not there is some unilateral benefit to the individual for cooperating, rather than all of the benefit being shared equally with others. If there is some direct benefit then public goods production or utilisation is a Snowdrift game, if not then it is a Prisoner's Dilemma. It should be stressed, though, that the Snowdrift game still represents a social dilemma (Sugden, 1986; Doebeli and Hauert, 2005), since the level of cooperation that is reached is still suboptimal in terms of mean fitness.

Table 1.1: A payoff matrix for social interactions (Doebeli and Hauert, 2005).

	Cooperate	**Selfish**
Payoff to Cooperate	$b - c/2$	$b - c$
Payoff to Selfish	b	0

A crucial assumption of the game theoretic approach is that it considers interactions in the population to be freely-mixed, as if the population was continuously being stirred so that who an individual interacts with at each time step is completely random, and not influenced by population structure. That is, the population is assumed to be unstructured with respect to social interactions. However, many social interactions will take place in groups that represent a subset of the population (Wilson, 1975a, 1980). We discuss in the next section how such interaction structure can allow greater levels of cooperation to evolve. In Chapter 5 we discuss the appropriateness of different types of game for modelling different kinds of ecological interaction, and for modelling the transitions to new levels of organisation in the biological hierarchy. We then demonstrate that the type of game affects the efficacy of the population structure-modifying process that we show in this dissertation to be an explanation for the origin of cooperation.

1.2 What is already known about how cooperation can evolve?

Much theoretical work in the last half century has examined how cooperative behaviours can evolve and be maintained in a population, despite their apparent individual cost (Hamilton, 1964a,b; Trivers, 1971; Wilson, 1975b,a; Axelrod and Hamilton, 1981; Maynard Smith, 1982; Frank, 1998; Hammerstein, 2003; Lehmann and Keller, 2006; West et al., 2007b; Fletcher and Doebeli, 2009). This work has shown that the answer lies in *interaction* structure[3]. Specifically, the population must be structured such that cooperative individuals tend to interact more frequently with each other than they would under free-mixing conditions. When this happens, the benefits of cooperation fall preferentially on other cooperators, and exploitation by cheating individuals is reduced. This can make cooperation the fitter behaviour, since cooperators enjoy the majority of the fitness benefits of cooperative behaviour, while cheats mainly interact with each other and hence forego this benefit. It can, therefore, be individually advantageous to cooperate, if doing so means that you will receive more of the benefits of others' cooperation yourself.

A vast body of literature (reviewed in Chapter 2) has explored the kinds of population that can give rise to such an interaction structure. The crucial aspects of population structure are those that affect an organism's interaction group – the subset of the population in which social interactions take place (Hamilton, 1975; Wilson, 1975a). Examples of aspects affecting cooperation include group size (Wilson and Colwell, 1981; Wilson, 1983a; Avilés, 1993, 2002), the degree of kinship between the interacting individuals (Hamilton, 1964a; Maynard Smith, 1964), and how frequently individuals disperse and interact with others outside of their local group (Taylor, 1992; West et al., 2002). All of these factors, and others, influence the degree to which the benefits of cooperation fall upon other cooperators.

Figure 1.1 shows an example population structure that makes cooperation advantageous, even when selfish cheats are present and can exploit cooperation. Individuals form groups and reproduce within them for a number of generations, t. Individuals then disperse from their groups, and mix freely in a migrant pool before founding the next generation of groups. Within groups, social interactions take place and affect individual fitness. Cooperative individuals can be exploited by selfish cheats within their group.

---

[3] This term is used by Michod and Sanderson (1985)

Thus, within each group, the proportion of cooperative individuals declines. However, after t generations the groups that were founded by a greater proportion of cooperators have grown to a larger size. Because of this differential group growth, cooperation can potentially increase in frequency globally when the groups are mixed and reformed. This is an instance of the statistical phenomenon known as Simpson's Paradox (Simpson, 1951), since the local decline in frequency of cooperators within each group does not preclude their global increase in frequency, since differential group productivity also has an effect (Sober and Wilson, 1998). The fact that the mechanism relies on differential group productivity means that it can be viewed as a process of natural selection acting between groups, i.e., group selection.

Crucial to this process is the presence of between-group variance: the groups must differ in their proportion of cooperators if some are to grow to a larger size than others, and hence if Simpson's Paradox is to operate (Sober and Wilson, 1998). Any mechanism that increases between-group variance will be more favourable to the evolution of cooperation (Wilson, 1975a; Hamilton, 1975). One such mechanism is kinship, for if the groups are composed of genealogical kin then they will tend to be more homogeneous, so group members will tend to be more similar to each other, but less like those from other groups (Hamilton, 1964a,b, 1975; Wade, 1980).

The particular population structure in Figure 1.1 is based on the classic "Haystack" model proposed by Maynard Smith (1964), and subsequently developed further by Wilson and Colwell (1981), Wilson (1987), and Sober and Wilson (1998). Such models are framed in terms of organisms that live on an ephemeral resource patch for a number of generations, before resource depletion triggers the dispersal stage and global mixing. It is often regarded as a classic example of a group selection model (Sober and Wilson, 1998). This population structure forms the basis of the models in this dissertation, and is discussed further in Chapter 3.

Population structuring, then, can allow cooperation to evolve. But we also need to understand how, exactly, to formulate the role of this structure in terms of a process of natural selection, for two reasons. The first is that from a pragmatic point of view, we need to know how to incorporate population structure into models of natural selection. For example if a population is subdivided into groups, how do we incorporate the effects of this group structure into the equations describing the dynamics of overall gene frequency change? The second reason is that how we describe the role of population structuring is fundamental to how we understand the process of natural selection itself. If, for example, a cooperative behaviour that benefits a whole group evolves, do we say this is an adaptation of the individuals, or of the group? Likewise, if that trait evolved because some groups had a greater proportion of cooperative individuals than others, did it evolve by selection at the level of the group, the individual, or even the gene? These issues have been much debated in both the biological literature (Williams, 1966; Dawkins, 1976; Wilson, 1975a, 1980; Hamilton, 1964a, 1975; Maynard Smith, 1987;

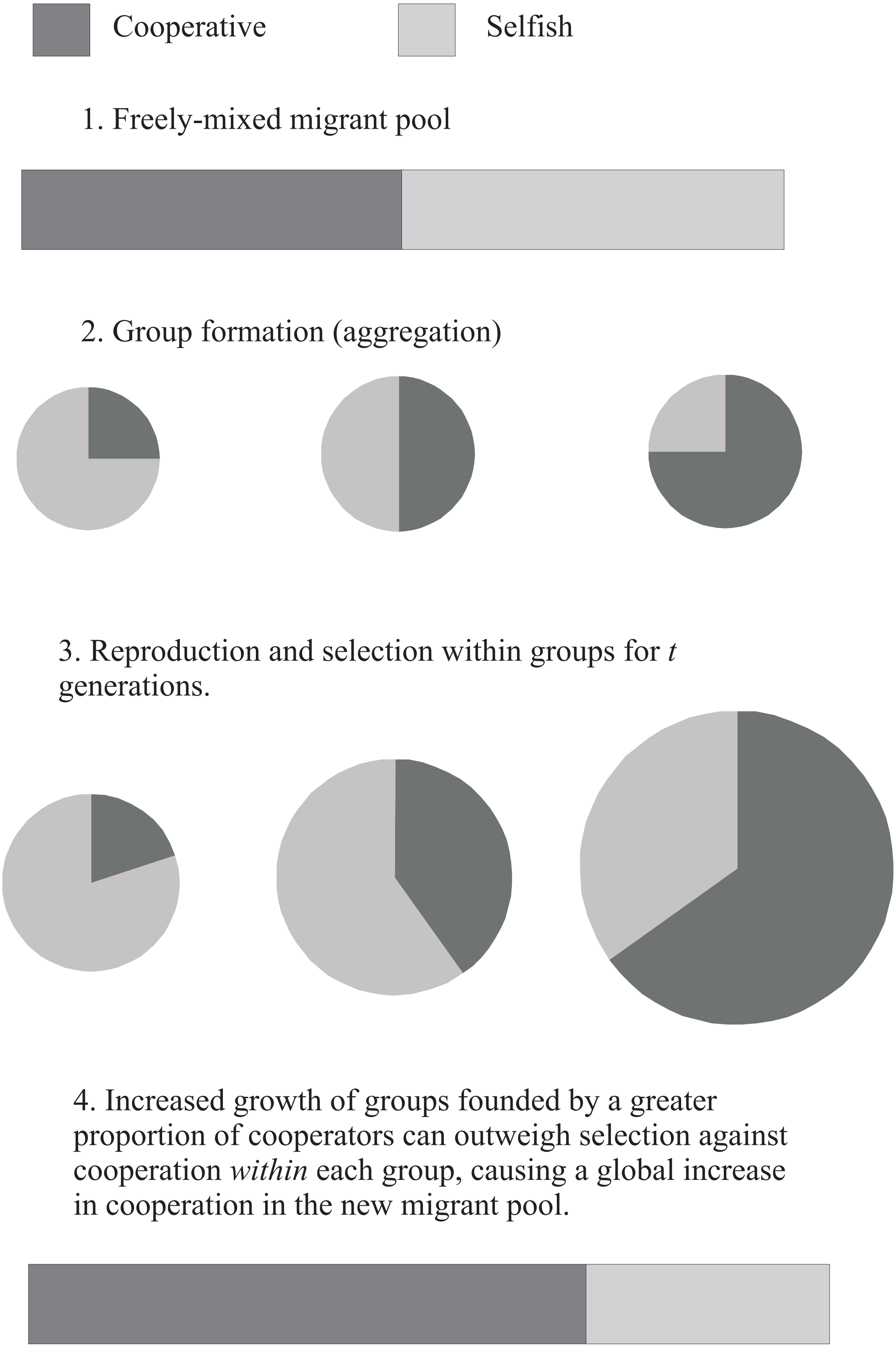


Figure 1.1: Cooperation can increase in frequency in the migrant pool due to differential group contributions, even though it decreases in frequency within each group.

West et al., 2007a) and the philosophical literature (Sober, 1984; Sterelny and Kitcher, 1988; Sterelny, 1996; Okasha, 2006), for many authors prefer to see natural selection acting only at the individual or genic level, while others are happy to ascribe processes of selection and adaptation to whole groups. This has led to different kinds of model, for example kin or group selection, providing different accounts of the role of the *same* population structure in selection. We address this issue in detail in Chapter 2, where we aim to resolve misunderstandings that are present in the literature, and argue that social niche construction helps to bridge the gap between group- and individual- selection.

1.2.1 The problem with our current understanding of how cooperation evolves

A plethora of models have shown the precise effects of various population structures on the evolution of cooperation (Williams and Williams, 1957; Hamilton, 1972, 1975; Wilson, 1975a, 1980, 1987, 1992; Wilson et al., 1992; Bull et al., 1991; Goodnight, 1992; Nowak and May, 1992; Doebeli et al., 2004; Fletcher and Zwick, 2004; Kreft, 2004; Peck, 2004; Kerr et al., 2006; Ross-Gillespie et al., 2007; Pepper, 2007). This means that for a given population structure, we can use this theory to predict how much cooperation can be supported. What these models can say very little about, however, is the *origin* of cooperation. That is, why is the population structured in such a way that selects for x amount of cooperation, rather than in a way that selects for y amount? If the structure was just a part of the selective environment, and could not change, then this would not matter. This is the simplifying assumption that most models of the evolution of cooperation make. However in reality population structure can evolve, for it is the product not only of environmental conditions, but also of individual behaviours. Many of these behaviours that affect population structure have a genetic basis, and so are themselves subject to natural selection. That this is the case is uncontroversial amongst evolutionary biologists, for the evolution of aspects such as the mating system (Orians, 1969; Emlen and Oring, 1977), group size (Rodman, 1981; Koenig, 1981), and dispersal rate (Johnson and Gaines, 1990) have all been considered. Despite this, however, there have been very few treatments that consider how their evolution is affected by selection pressures on (individually-costly) cooperative behaviour. A very small number of works do consider this in specific cases, for example, Peck and Feldman (1988),Breden and Wade (1991), and Avilés (2002). Here we generalise such works, and show how the creation of selection pressures on social behaviour is a fundamental and overlooked driving force in the evolution of population structure.

1.3 Cooperative behaviour originates from social niche construction

We develop here a theory that explains why organisms live in a social environment that supports a certain level of cooperation. To do so, we consider the joint equilibrium of population structure and social behaviour, by allowing both to be influenced by heritable genetic traits and hence to evolve in an adaptive manner. It is well appreciated that genetic traits can affect the pertinent aspects of population structure, just as other genetic traits can affect social behaviour. Examples of genetic traits that modify population structure include those affecting group size (McCauley, 1994), the degree of kin discrimination in social interactions (Griffin and West, 2003), and the dispersal rate (Olivieri et al., 1995). Such traits exist across all taxa; for example, bacteria in a biofilm may be able to influence the size of their micro-colony through the amount of extracellular polymeric substances that they secrete (Flemming et al., 2007), plants can evolve the dispersal distance of their seeds (Levin et al., 2003), and some fish can choose with whom they interact with respect to certain social behaviours such as foraging or predator avoidance (Dugatkin and Wilson, 1993; Griffiths, 2003). Moreover, such changes in population structure directly affect the selection pressures on social behaviour that their bearers experience. For example, the creation of a cell membrane by the early replicators locked them into a group selection process that selected against competitive exclusion, and allowed for coexistence of different types, as shown in the Stochastic Corrector model (Szathmáry and Demeter, 1987). Likewise, the evolution of vertical transmission of symbionts, as in eukaryotic cells and their organells (Margulis, 1970), selects for cooperation between host and symbiont.

Niche construction is a term for the process by which individuals modify their environment, and in so doing change the selection pressures that they experience on other traits (Odling-Smee et al., 2003; Laland and Sterelny, 2006). In this way organisms can actively create their own selection pressures (Lewontin, 1982, 1983), rather than simply being the passive recipients of those brought about by environmental conditions. We argue here that a type of niche construction process can operate on population structure, whereby the evolving individuals create the conditions for more or less cooperation. Most models of social evolution assume that the population structure of the evolving organisms is a static aspect of the environment, but the niche construction paradigm suggests that this can in fact change, and will do so in an adaptive manner. That is, organisms create their own social niches in which either cooperation or selfishness thrive.

This concept of *social niche construction* is illustrated in Box 1.1. Conventional models of social evolution show the equilibrium frequency of cooperation that is reached under a given population structure. But given that population structure can itself evolve, the fundamental question is: what is the equilibrium that is reached in the combined space of structures and behaviours? That is, we must explain why organisms live in a population structure that creates a certain equilibrium frequency of cooperation. Classical models do not consider this, for they make a simplifying assumption which amounts to saying that the structural equilibrium is provided by fiat, and not in any way affected by the opportunity for more or less cooperation to evolve. Thus, these models provide an unsatisfactory explanation for the true origin of cooperation, as they cannot explain why organisms live in a structure that supports a certain amount of social behaviour.

It should be noted that the term "social niche construction" has also been used in a more limited sense to refer to the construction of social networks by evolving individuals, particularly in humans and primates (see, for example, Flack et al. 2006). This work typically assumes cognitive abilities on the parts of the organisms, and considers how direct relationships between individuals can evolve. However we argue that in general, social niche construction can be much simpler, and does not require cognitive abilities. Rather, it just requires that organisms can modify their population structure, which applies even at the origins of life, for example, with the creation of a cell membrane by the early replicators (Szathmáry, 1999). Thus, organisms affecting their social niche can be a much more general dynamic than studies focussing on higher organisms have suggested.

We demonstrate here that social niche construction can drive the evolution of cooperation and explain its origin. Specifically, we show that organisms will actively modify their population structure to *create* social selection pressures favouring greater cooperation. Furthermore, we will argue that the direction of any component of selection on population structure, that is due to social behaviour, must be towards a structure that supports cooperation, rather than selfishness.

Niche construction = organisms modify their environment and hence the selection pressures they and future generations experience.

+

Social evolution = the evolution of cooperative or selfish behaviours as determined by the population structure.

Social niche construction = Organisms modify their population structure and hence the selection pressures they and future generations experience on the evolution of social traits.

Box 1.1: Social niche construction explains the equilibrium of structure and social behaviour that a population reaches.

We illustrate this argument by constructing a model for the origin of both cooperation, and the population structures that support it. One trait in our model controls social behaviour (where we take a social trait to mean behaving either cooperatively or selfishly towards other group members) as in conventional models of social evolution, while another influences an aspect of population structure, such as the initial size of the bearer's group. In this model we have shown how a population can start out with either no or little cooperation between its members at the initial behavioural equilibrium, but can, through the evolution of population structure, end up in an equilibrium state where cooperation dominates.

It should be noted that whilst we consider two separate traits, one of which affects social behaviour and the other population structure, it is the case that the same trait could affect both of these. For example, extracellular matrix production by bacteria in a biofilm affects population structure, but is also a social trait since the extracellular matrix is a public good that is costly to produce. Thus, although we consider separate traits to illustrate our argument, the consideration of single traits that affect both social behaviour and population structure is an important avenue for future research.

1.4 Claims of the thesis

In this dissertation, we provide evidence for the following claims:

- We show that social behaviour can exert selection pressure on population structure-modifying traits, causing individuals to adaptively modify their population structure to support greater cooperation. We investigate conditions under which this occurs, and find the dynamic to operate from a much larger range of initial conditions when selection on a socio-behavioural trait supports a protected polymorphism of behaviours, as in the Snowdrift game.
- Any component of selection on structure-modifying traits that is due to social behaviour must be in the direction of increased cooperation, rather than increased selfishness.

1.5 Logical argument for social niche construction

The logical argument in Box 1.2 shows that any component of selection on population structure due to social behaviour must be in the direction of that which supports increased cooperation.

Our argument states that if individuals have a trait that causes them to preferentially live in a certain structure, and that structure exerts a different selection pressure on social behaviour, then linkage disequilibrium must evolve between the socio-behavioural and population structuring-traits. The mean fitness of individuals in a structure that selects for cooperation will be greater, all other factors being equal, and this direct selection on social behaviour drives indirect selection on the structure-modifying trait. Thus, if structure A creates selection pressures on social behaviour favouring more cooperation, and structure B creates selection pressures favouring more selfish behaviour, then the

Definition Cooperation increases the mean fitness of population members (Wilson, 1975a; Dawes, 1980).

Premises

1. Population structure affects selection pressure on social behaviour, as proved by Hamilton (1964a, 1975) and Price (1972).
2. Structure A selects for greater cooperation, structure B for greater selfish behaviour (made logically possible by premise 1).
3. Individuals have a heritable trait that produces population structure A or B. Individuals also posses a heritable social behaviour (cooperative or selfish).

Assumption Individuals with trait A tend to find themselves in structure A more frequently than individuals that do not posses the trait. Likewise with trait and structure B.

Argument

1. Linkage disequilibrium evolves between behavioural and structural traits, since the cooperative trait increases in frequency more in structure A than structure B (premise 2), and the individuals in structure A are more likely to have the A structural trait (due to the assumption above). Conversely, the selfish trait increases in frequency more in structure B by the same logic.
2. Since by definition cooperation increases mean fitness, and cooperation has become linked with the A structural trait, individuals with the A structural trait have the component of their fitness that is due to social behaviour (i.e., fitness affects from interactions with others) increased, on average, compared to individuals with the B trait.

Box 1.2: Logical argument that any component of selection pressure on population structure, that exists due to social behaviour, must be towards a structure that increases cooperation.

component of individual fitness that comes from social behaviour will be increased by more in the individuals that live in structure A. If these are the same individuals that have the A structure-creating trait (assumption in Box 1.2), then that structure-creating trait will be fitter than its rival allele, assuming all other selective forces remain the same across both structures.

This argument relies on the assumption in Box 1.2, that individuals with trait A tend to find themselves in structure A more frequently than individuals that do not posses the trait. Without this assumption, linkage disequilibrium would not be generated between the socio-behavioural and structure-modifying traits. This is because although individuals that happened to live in structure A would be fitter, due to enjoying a greater frequency of cooperation, this increase in fitness would not be linked with the A trait, but would fall upon both structure-modifying traits equally. Essentially, there needs to be a structure differential between individuals with the A trait and individuals with the B trait. This then entails assortativity on the structure-modifying trait, such that individuals with a particular structural trait tend to be segregated to some degree from the rest of the population. This can clearly occur if individuals actively modify their own structure, for example, by changing their habitat preference. What happens, however, if the trait changes the structure of other individuals, but does not affect the structure of self? A structure differential would still be created in this case, and the A trait would still be selectively advantageous if it made other individuals *less* cooperative. That is, what matters is the frequency of cooperation experienced by bearers of the A trait *relative* to non-bearers. So any trait that increases the relative frequency of cooperation in its bearers will be favoured by a component of selection, even if that relative increase is achieved by decreasing the level of cooperation in non-bearers. On the other hand, if the trait made other individuals live in a structure that was more cooperative, then it would be selected against, because it would increase the fitness of non-bearers. A component of selection on a structure-modifying trait, that is due to social behaviour, must therefore increase the level of cooperation in its bearers relative to non-bearers. The fact that there is a differential in cooperation between bearers and non-bearers of the structure-modifying trait allows linkage disequilibrium to evolve, and hence the structure-modifying trait to be selectively advantageous due to it becoming linked with relatively more cooperation.

The evolution of population structures that support cooperation by this mechanism takes the form of a positive feedback process between selection for cooperation, and selection for the structure that supports cooperation. Specifically, if greater cooperation is selected for in a structure, then the individuals that have a trait which creates that structure have a component of their fitness increased. Thus, that structure-creating trait increases in frequency in the population, so that even more individuals will then create and live in the structure, which in turn selects for even more cooperation globally. Hence, cooperation selects for the population structures that support it, and selection for these structures creates a selective environment that favours even more cooperation. A similar positive feedback argument was made by Breden and Wade (1991) for the specific case of inbreeding evolving concurrently with cooperation. They termed such a positive feedback process "Runaway Social Selection". Our thesis builds on such work by constructing a general logical argument that is not specific to mating structure, and investigating the effects of other (direct) components of selection on structuring traits that can oppose those due to social behaviour. We thus generalise works such as this and Peck and Feldman (1988), which considered the concurrent evolution of monogamy of social behaviour, and draw out the general implications of such processes for social evolution through our social niche construction hypothesis. In particular, it is argued here that such processes are fundamental to explaining the evolutionary origin of cooperation, and in providing an adaptive explanation for the creation of new levels of individuality, as in the major transitions to new levels of biological organisation (Maynard Smith and Szathmáry, 1995).

1.6 What are the factors affecting whether social niche construction will occur?

The logical argument presented above will hold if the premises are true. The first of these, that population structure affects selection pressure on social behaviour, has been proved by Hamilton (1964a) and Price (1972). We argue that the third, that individuals posses heritable traits affecting population structure, is also likely to hold in most organisms across all taxa, as discussed in Section 1.3. The remaining premise in the social niche construction argument is that structure A induces selection for greater cooperation, and structure B conversely selects for greater selfish behaviour. If new structures arise as a result of mutations, then structure B could just be the current structure, which selects for greater selfishness relative to the new mutant structure. For structure A to select for greater cooperation, it must increase the assortment of cooperators (Hamilton, 1964a), or equivalently, the between-group variance in social behaviour (Wilson, 1975a; Hamilton, 1975), by a sufficient amount to change the selection pressure on the socio-behavioural trait. If the new structure arises by genetic mutation, the difference between it and the current structure can be expected to be small. Thus, in order for the premise to be satisfied, a small change in structure needs to increase between-group variance by a sufficient amount for greater cooperation to be selected for. We have investigated in detail one factor that affects this, which is whether a polymorphism of social behaviours is supported within groups, as in the Snowdrift game. It is shown here in chapters 4 and 5 that such a polymorphism, as can result from negative frequency-dependent selection on social behaviour, greatly increases the range of conditions over which a small change in structure increases cooperation. Essentially, this is because the amount of between-group variance that is possible depends on the frequency of the least common type in the population (Wilson, 1980, 1983a), and this can tend to zero over a large range of parameter space under directional selection. If there is a polymorphism of behaviours, however, then the possibility of between-group variance being generated is preserved, so a small change in structure can lead to an increase (Powers et al., 2008).

Our logical argument concludes that any *component* of selection on population structure due to social behaviour must be in the direction of increased cooperation. There can, however, be other components of selection on structural traits that are in opposition to this. The direction in which population structure evolves, if at all, then depends on the balance of these different forces. We investigate this by, as an example, modelling the concurrent evolution of initial group size preference and social behaviour, as described in chapters 3 and 6. In these models, smaller initial group size supports greater cooperation, through greater between-group variance, but a larger initial group size provides intrinsic benefits due to an Allee effect (Allee, 1938; Avilés, 1999).

1.7 Original contributions presented in this dissertation

The broad original contributions produced in the development of the thesis are as follows:

- The argument that social niche construction provides an explanation for the origin of cooperation, by showing how the population structures that select for cooperation evolve as a result of individual adaptation.
- The argument that any component of selection on population structure-modifying traits, that is due to social behaviour, must be in the direction of a structure that selects for greater cooperation.
- The argument that social niche construction can help to resolve the tension that exists between group- and individual- selectionists in the social evolution literature. Individualists often see group selection as just a case of individuals adapting their social behaviour to the group structure that they find themselves in. That is, the groups are treated as part of the context for individual selection, like any other environmental factor. However, when individuals can modify their group structure, this changes the selection pressures they experience on social behaviour. In that case, we cannot view group selection as just individuals adapting their behaviour to their social environment.
- An examination of the conditions under which selection on population structure-modifying traits, due to social behaviour, occurs. If social niche construction is to occur by small mutations on structure-modifying traits, then a small change in population structure must result in an increase in between-group variance. We show that this occurs from a much larger range of initial conditions where a protected polymorphism of social behaviours occurs within groups, for example, as a result of negative frequency-dependent selection.
- The argument that social niche construction can help to provide a systematic adaptive explanation for some of the major transitions in evolution.
- A new understanding of the importance and role of different types of evolutionary games when considering social evolution in general, and the major transitions in particular.

The following original contributions were also developed during the supporting modelling work:

- Demonstration that when social behaviour and population structure-modifying traits are concurrently subject to selection, non-monotonic dynamics on the allele frequencies of both traits can arise. In particular, selfish behaviour, and the structure-modifying traits that support it, can initially increase in frequency, yet cooperative behaviour and its supporting structure can eventually fix in the population.

- The understanding that where within-group selection supports a polymorphism of behaviours (e.g., due to negative frequency-dependent selection), then some effect of group selection can be seen over a much larger range of parameters. We find that this also holds even under assortative group formation, hence is relevant to traditional kin selection models.
- A model comparing the efficacy of group selection when cooperation takes the form of a Snowdrift rather than Prisoner's Dilemma game, illustrating the above understanding.

1.8 Published research contributions

The following first author publications have been produced during the period of candidature:

- Powers, S. T., Penn, A. S., and Watson, R. A. (2007). Individual selection for cooperative group formation. In Almeida e Costa, F., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life*, volume 4648/2007 of *Lecture Notes in Computer Science*, pages 585594. Springer Berlin / Heidelberg.
- Powers, S. T. and Watson, R. A. (2007). Preliminary investigations into the evolution of cooperative strategies in a minimally spatial model (poster). In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, page 343. ACM Press.
- Powers, S. T. and Watson, R. A. (2007). Investigating the evolution of cooperative behaviour in a minimally spatial model. In Almeida e Costa, F., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life*, volume 4648/2007 of *Lecture Notes in Computer Science*, pages 605614. Springer Berlin / Heidelberg.
- Powers, S. T., Penn, A. S., and Watson, R. A. (2008). The efficacy of group selection is increased by coexistence dynamics within groups. In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 498505, Cambridge, MA. MIT Press.

- Powers, S. T. and Watson, R. A. (2008). The group selection debate and ALife: Weak altruism, strong altruism, and inclusive fitness (abstract). In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, page 796, Cambridge, MA. MIT Press.
- Powers, S. T. and Watson, R. A. (2009). Evolution of individual group size preferences can increase group-level selection and cooperation. In *Advances in Artificial Life: Proceedings of the 10th European Conference on Artificial Life*, volume 5777, 5778 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg.
- Powers, S. T., Mills, R., Penn, A. S., and Watson, R. A. (2009). Social environment construction provides an adaptive explanation for new levels of individuality. In *Proceedings of the ECAL 2009 Workshop on Levels of Selection and Individuality in Evolution: Conceptual Issues and the Role of Artificial Life Models.*
- Powers, S. T., Penn, A. S., and Watson, R. A. (in prep.). The concurrent evolution of cooperation and the population structures that support it.

The following contributing author publications were also produced:

- Penn, A. S., Powers, S. T., Conibear, T., Kraaijeveld, A., Watson, R., Bigg, Z. and Webb, J. (2008). Co-operation and Group structure in Bacterial Biofilms. In: *Society for General Microbiology, Autumn meeting.* Trinity College, Dublin.
- Watson, R. A., Mills, R., Penn, A. S. and Powers, S. T. (2008) Can individual selection favour significant higher-level selection? (abstract). In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, page 818, Cambridge, MA. MIT Press.
- Penn, A. S., Watson, R., Powers, S. T., Webb, J., Kraaijeveld, A., Conibear, T. and Bigg, Z. (2008). Mechanisms for the initiation of multicellularity in bacterial biofilms (abstract). In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, page 794, Cambridge, MA. MIT Press.
- Snowdon, J. R., Powers, S. T. and Watson, R. A. (2009). Moderate contact between sub-populations promotes evolved assortativity enabling group selection. In *Advances in Artificial Life: Proceedings of the 10th European Conference on Artificial Life*, volume 5777, 5778 of *Lecture Notes in Computer Science.* Springer Berlin / Heidelberg.

- Watson, R. A., Palmius, N., Mills, R., Powers, S. T. and Penn, A. S. (2009). Can Selfish Symbioses Effect Higher-level Selection?. In *Advances in Artificial Life: Proceedings of the 10th European Conference on Artificial Life*, volume 5777, 5778 of *Lecture Notes in Computer Science.* Springer Berlin / Heidelberg.

1.9 The implications of the thesis for our understanding of cooperation and group selection

It is already widely appreciated that it can, in certain situations, be individually advantageous for an organism to cooperate, even if this comes at an apparent direct fitness cost. What has largely been left unaddressed, however, is why such situations exist in nature. One answer is that they come purely from the environment, and hence do not need an evolutionary explanation. We would argue that this is an implicit assumption in most social evolution models, for aspects of population structure that directly affect social evolution are assumed to be static. However given that organisms can modify their environment, then surely we should consider evolution operating on environment as well as behaviour. Using the arguments developed here, we can explain how organisms come to find themselves in a situation where it pays to cooperate: they create the conditions favouring increased cooperation themselves, because this increases their fitness compared to individuals that do not create those conditions. Our argument implies that feedback between environment and behaviour can be a fundamental driver in social evolution, which we have demonstrated in a model of the concurrent evolution of initial group size and social behaviour.

The argument that we develop is fully compatible with the “selfish gene” view of evolution (Dawkins, 1976). This is because at all times in our argument, population structure changes because the individuals who change it are fitter than those that do not, hence the allele for changing the structure is fitter in the population. Thus, we can provide a systematic account fully compatible with natural selection. However, the fact that we show how organisms themselves create the conditions for cooperation casts new light on what we really mean by “cooperation” and “selfishness”. In particular, what is the cooperation relative to? When the environment is fixed, the individuals are simply doing the social behaviour that increases their fitness in that particular environment. It is thus hard to see such behaviours as anything but selfish, regardless of whether they end up also helping others. One only labels behaviours as cooperative in this context because they help others more than they would do so in a freely-mixed population. But, this freely-mixed situation is entirely hypothetical. The reality is that individuals live in a certain population structure and selfishly help each other more or less depending on how much population structure there is.

To illustrate this, consider a typical model for the evolution of social behaviour in a structured population, such as that introduced previously in Figure 1.1. In such models, the authors show conditions under which a rare mutant cooperative allele can increase in frequency. The conclusion is then that they have shown the evolution of cooperative behaviour. However, they have not shown an *increase* in equilibrium cooperation, because the frequency at which they started out the cooperative allele was never stable under selection in that population. Rather, the models typically start out close to the frequency that would be stable in a hypothetical freely-mixed population (i.e., close to zero). Thus any evolution of cooperation in such a model is relative only to that which would occur in a hypothetical world. In particular, the cooperative allele is started out of equilibrium for the population structure that is present.

When we consider that the population structure can itself evolve, however, then matters are different. This is because we do then have a meaningful baseline case to compare the level of cooperation that the population evolves to - it is the initial population structure that previous generations of the organisms actually *did* live in. Thus, we are not comparing the degree of cooperation present in the current population structure with a hypothetical case, we can compare it with the equilibrium level of cooperation in the structures that the organisms used to live in. We therefore have an actual initial level of cooperation to compare any subsequent evolution to. In that sense, we *can* say that organisms are evolving to become more cooperative, rather than just adapting to their current environment, for we start the cooperative allele *in* equilibrium for the current environment. Of course, when organisms create the conditions for greater cooperation, they are doing so because it selfishly increases their (or their genes) fitness. But, when considering social behaviour, we can see a meaningful increase in cooperation because the equilibrium frequency of cooperation changes over time. Thus, the evolution of cooperation is not just a cooperative allele reaching its equilibrium frequency, as in standard models. Rather, we show how feedback between changes in environment and changes in social behaviour move the equilibrium. This can, of course, be viewed as population structure being started out of equilibrium, and social behaviour and population structure then moving to their joint equilibrium. However, while the vast majority of models assume that structure starts in equilibrium, our thesis explains how this equilibrium is reached by individual adaptation. We are thus able to address a question that most models cannot: how do the conditions that support cooperation arise? Whereas previous approaches view the origin of cooperation as a mutant cooperative allele being selected for, we argue that the only satisfactory explanation for the origin of cooperation is one that explains the evolution of an interaction group structure that selects for cooperative behaviour.

This argument bears on the role of groups in group selection and kin selection models. In models where the group structure does not itself evolve, it is easy to sideline the role of groups. In particular, it is easy to treat groups as just part of the context in which individual selection occurs (Dugatkin and Reeve, 1994; Sterelny, 1996; Kerr and Godfrey-Smith, 2002), since the individuals are just adapting their behaviour to the group structure that exists, why invoke notions of group selection? The philosophy of individualists (e.g. Foster et al. 2006a; West et al. 2007a, 2008) is implicitly that it is a fact of life that some organisms live in groups, and that they adapt their behaviour accordingly in order to selfishly increase their individual fitness. For example, inclusive fitness equations (Hamilton, 1964a) recognise group structure, but the groups are only implicitly defined by the coefficient of relatedness between actor and recipient of cooperation. But when considering the creation of this group structure, it is not so easy to sideline the role of groups in the selective process. If the organisms are themselves creating the groups then surely the groups cannot be seen as just part of the selective context, for that context is itself changing in an adaptive manner. Nearly all group selection models just consider how social behaviour adapts to the current group structure. That is, they show how an equilibrium frequency of a social allele, under the selection pressure from the current group structure, is reached. By contrast, in this dissertation we show how a group-level selection pressure on social behaviour is created, making any explanation that views groups as being something that organisms merely adapt their social behaviour to much less satisfactory.

1.10 The implications of social niche construction for the major transitions in evolution

Table 1.2 lists the events in the history of life on earth that Maynard Smith and Szathmáry (1995) mark out as the “major transitions in evolution”. Essentially, several of these transitions involved formerly free-living individuals aggregating into cooperative groups. For example, replicating molecules aggregated into compartments, independent replicators joined together to form chromosomes, and solitary insects aggregated into colonies. A striking feature of these transitions is that the groups became reproductively inseparable, such that their members became only able to reproduce by reproducing the group as a whole (Maynard Smith and Szathmáry, 1995). Such a reproductive strategy is a pinnacle example of cooperation, in particular, of strong altruism.

The transitions can be conceptualised very naturally in terms of multi-level selection (Buss, 1988; Maynard Smith and Szathmáry, 1995; Michod, 1999). This is because selection *within* the aggregations will favour cheating behaviour, for example, individuals that do not forego their own reproduction in order to reproduce the group as a whole. However, selection *between* groups will favour those with greater cooperation amongst their members. Moreover, a transition could be said to occur when within-group selection becomes (almost) completely suppressed, leaving between-group selection as the driving force in evolution. For example, it is between cell-group selection, and the virtual

Pre-transition	Post-transition
Replicating molecules	Populations of molecules in compartments
Independent replicators	Chromosomes
RNA as gene and enzyme	DNA + protein
Prokaryotes	Eukaryotes
Asexual clones	Sexual populations
Protists	Animals, plants, fungi (cell differentiation)
Solitary individuals	Colonies (non-reproductive castes)
Primate societies	Human societies (language)

Table 1.2: The major transitions in evolution (as defined in Maynard Smith and Szathmáry 1995, p.6).

absence of within-group selection, that allows multicellular organisms to evolve adaptations and hence be evolutionary individuals in their own right (Michod, 1999). That is, the complete dominance of between- over within-group selection allows the groups to be "units of evolution", *sensu* Maynard Smith (1987). How between-group selection comes to dominate within-group selection is, therefore, something that an adaptive theory for the major transitions must address.

The transitions can also be further classified into what Queller (1997) calls fraternal and egalitarian[4]. In the fraternal transitions, the pre-transitional entities are of the same type, and so there is the potential for a reproductive division of labour in which one individual can reproduce the whole group (Queller, 1997). In particular, kinship amongst the cells can provide a high between-group variance and eliminate within-group variance, as occurs when multicellular organisms develop from a single zygote, for example (Maynard Smith and Szathmáry, 1995; Grosberg and Strathmann, 1998; Queller, 2000). On the other hand, in the egalitarian transitions the pre-transitional entities are of different species, hence a reproductive division of labour cannot occur because the pre-transitional individuals do not share genes (Queller, 1997). However, the fact that they are of different types means that there can be a combination of different functions from the outset (Queller, 1997), for the types will naturally perform different functions. Thus, while division of labour must be evolved in the fraternal transitions, the pre-transitional entities already performed different (complementary) functions in the egalitarian transitions. For example, it has been suggested that different types of replicator provided different but essential contributions to a common metabolism in the first protocells (Szathmáry, 1999). So, the initial advantage in forming groups in the egalitarian transitions may have been a combination of functions, while in the fraternal transitions such a division of labour would have to be evolved and so the initial advantage must have been something else, such as economies of scale through larger size (Queller, 1997).

---

[4] Although some transitions do not fit into the fraternal / egalitarian distinction, for example, the origin of the genetic code.

Explaining the major transitions essentially involves explaining the evolutionary origin of highly cooperative groups. That is, how and why did the individuals change their population structure to create a selective environment favouring cooperation? This is the very problem that our thesis of social niche construction addresses. In particular, we are able to show how individuals evolve to create groups that exhibit ever greater degrees of cooperation, for the benefits of greater cooperation can drive the evolution of population structure and vice versa. Crucially, this happens gradually by selection on a series of small mutations. As a result, we can provide a systematic explanation for the creation of some of the prerequisite conditions for a transition that relies only on the selection of small mutations. This gives support to the notion that there is a common and systematic selective process behind several of the major transitions, such that while the details of any one transition are contingent on natural history[5], there are unifying selective forces at work across transitions. It is worth noting that transitions can sometimes be reversed[6], for example, some sexual organisms also facultatively reproduce asexually by parthenogenesis, and multicellularity is thought to have evolved and then been reversed back to unicellularity in some taxa (King, 2004). When this happens, we suggest that other components of selection pressure on population structure, apart from those resulting from social behaviour, must become stronger, such that the benefits of cooperative group living are no longer strong enough to keep the group together. That is, ecological factors may change such that they weaken the benefits of cooperation, and hence weaken the component of selection pressure on population-structure modifying traits that derives from cooperation. When this happens, population structure may evolve in a direction opposite to that which supports cooperation, thereby undoing the transition. Thus, both the creation and reversal of transitions fits into the social niche construction framework.

Our model for the evolution of initial group size is particularly relevant to some of the fraternal transitions (*sensu* Queller). In particular, the transitions from single- to multi- celled organisms, and from solitary insects to eusocial colonies. In each of these cases, the initial group size is low: most multicellular organisms develop from a single fertilised egg, while eusocial insect colonies develop from one or a few queens. This raises the question: if there is advantage in numbers, as there must be if the evolution of cell-groups and insect colonies was adaptive, then why is the population structure such that they periodically return to one or a few cells or queens (Maynard Smith and Szathmáry, 1995; Grosberg and Strathmann, 1998; Roze and Michod, 2001; Wolpert and Szathmáry, 2002)? It has been hypothesised that this occurs in order to increase cooperation between members of the cell-group or colony, for if development is from a

---

[5] Transitions can be contingent either on the necessary genetic variation being present in the population (variation limited, e.g., the origin of the eukaryotic cell), or on the presence of an appropriate ecological context for that variation to be selected (e.g., the origin of multi-cellular organisms; King 2004) (Szamado and Szathmary, 2006).

[6] Some of the transitions happened multiple times independently, e.g., multicellularity, whereas other happened only once, e.g., the origin of the eukaryotic cell. The transitions that happened only once, which were therefore hard for selection to recreate, are interestingly the ones that have not been reversed.

small number of founders then cooperators are more likely to be clustered together. In other words, the variance on social behaviour within groups will be decreased, while that between groups will be increased, leading to a statistical clustering of social behaviours. Such clustering is, as we know, essential for the evolution of cooperation. Our thesis can explicitly explain how this clustering itself evolves. Moreover, on some accounts of natural history, initial group size gradually decreased as in the models presented here. For example, it has been suggested that the multicellular ancestors of modern day metazoans reproduced not by sending out single cells but rather by sending out fragments of many cells (Michod and Roze, 2000). Indeed, some loose examples of multicellularity do not develop from a single cell at all, but form by the aggregation of many cells, as is the case with cellular slime moulds. The "multicellular" forms that do this appear less well functionally integrated, and show greater cheating behaviour (Pál and Papp, 2000). The advantages of suppressing cheating may therefore create selection pressure for a bottleneck during development, whereby a cell-group or insect colony is founded by one or a few individuals (Maynard Smith and Szathmáry, 1995; Michod and Roze, 2000; Roze and Michod, 2001; Queller, 2000; Godfrey-Smith, 2009). Our thesis suggests how such a bottleneck can arise adaptively and gradually, by selection on a series of mutations that control initial group size and hence the degree of the developmental bottleneck.

1.11 Outline of dissertation

The remainder of this dissertation is organised as follows:

- Chapter 2 considers the tension between group selection and inclusive fitness perspectives on social evolution, that has been the source of much controversy in the literature since the 1960s. We consider arguments for and against regarding the role of groups in evolution as being merely the context in which individual selection occurs. We conclude that social niche construction posits a much more significant role for groups in social evolution than authors with an individualist philosophy have previously suggested. Part of this argument was presented in a talk at the 11th International Conference on the Simulation and Synthesis of Living Systems (Powers and Watson, 2008).
- Chapter 3 presents a minimal model of social niche construction, where a trait of cooperative or selfish behaviour evolves alongside a trait that produces a population structure where groups are either founded by many (low between-group variance) or few (high between-group variance) individuals. We show that smaller groups, which select for greater cooperation, evolve. We find conditions under which this occurs despite there being an opposing component of selection on population structure due to an Allee effect. The model provides a minimal illustration

of the logical argument in Box 1.2. The model from this chapter is presented in Powers et al. (2007).

- Chapter 4 investigates the effect of negative frequency-dependent selection on social behaviour, with respect to the range of conditions over which an effect of group selection can be seen. We find that a stable coexistence of social types, modelled by the Lotka-Volterra competition equations (May, 1976), greatly increases the range of group sizes over which some effect of group selection can be seen, and hence the range over which a small change in group size selects for greater cooperation. This is important for social niche construction, since selection on population structure, that is due to social behaviour, requires that a change in structure induces a change in social selection pressure (premise 2 of the logical argument in Box 1.2). Our results suggest that this is more likely to occur under coexistence dynamics. This chapter is based on material in Powers et al. (2008).

- Chapter 5 argues that the Prisoner's Dilemma game, which represents directional selection favouring the most selfish type, has dominated the social evolution literature, despite the fact that many types of biological social interaction are not best modelled by this game. We discuss 4 classes of game, and introduce a scale as to how conducive they are to group selection. We also discuss the appropriateness of different types of game for modelling different kinds of major evolutionary transition, particularly regarding the fraternal versus egalitarian split of the transitions proposed by Queller (1997). We conclude that types of game apart from the Prisoner's Dilemma have been unduly under-considered, are more conducive to be acted on by group selection, and are a more appropriate model for some of the transitions.

- Chapter 6 builds on the illustrative model of social niche construction presented in Chapter 3 by considering the introduction of new group size preference alleles by mutation, rather than all possible genotypes being present from the outset. This model illustrates how selection on a series of small mutations on population structure can create the conditions for cooperation. We find that this is more likely to occur where selection on social behaviour is negative frequency-dependent, as predicted by our results from Chapter 4. The model from this chapter, and parts of the analysis, are presented in Powers and Watson (2009).

- Chapter 7 argues that social niche construction is fundamental to providing an adaptive explanation for many of the major transitions. We discuss how whether a transition is of the fraternal or egalitarian kind constrains the type of social niche construction that can occur. Part of this argument is based on material presented in Powers et al. (2009).

- Chapter 8 reiterates the social niche construction hypothesis and its importance, and discusses the advantages and limitations of the modelling approach that we have used to illustrate it. This chapter also discusses how social niche construction can be applied in various other contexts apart from those considered in this dissertation.

Chapter 2

The Group Selection Debate and Social Niche Construction

In this chapter, we discuss the group selection debate that has raged in the social evolution literature since the 1960s, and how the social niche construction framework can help to resolve some of the tensions that are present.

2.1 Outline of our argument

Group selection has been much debated in the evolutionary biology literature since the 1960s, and remains highly controversial to this day (Wynne-Edwards, 1962; Maynard Smith, 1964, 1976, 1987; Williams, 1966; Wilson, 1980, 2007; Wilson and Sober, 1994; Wilson and Wilson, 2007; Sober and Wilson, 1998; Queller, 2004; Foster et al., 2006a; Okasha, 2006; West et al., 2007a, 2008). On the face of it, the concept of group selection appears simple: natural selection can logically act between groups to cause the spread of a group-beneficial trait, just like it can act between individuals to cause the spread of an individually-beneficial trait (Sober and Wilson, 1998). All parties agree on this point: no one has argued that such selection is *logically* impossible. However, many have vigorously suggested that such selection is unlikely to be important in nature, for they argue that there will rarely be sufficient variation in the (genetic) composition of groups for selection amongst them to be effective (Maynard Smith, 1964, 1976; Williams, 1966). In particular, group selection is often considered to be a highly dubious explanation for the evolution of cooperation, since it is suggested that individual selection will sweep selfish behaviour to fixation within each group, destroying any between-group variance (Williams, 1966; Maynard Smith, 1976; Dawkins, 1976). This conclusion was supported by the early models of group selection, in which the groups were demes whose probability of extinction was inversely proportion to their frequency of cooperative behaviour, and which reproduced by producing migrants that colonised vacant patches (Boorman and Levitt, 1972; Levin and Kilmer, 1974).

David Sloan Wilson, however, has led a challenge to this position, by arguing that the conception of groups as reproductively isolated demes is unnecessarily narrow (Wilson, 1975a, 1977, 1980; Wilson and Sober, 1994; Sober and Wilson, 1998). Rather, alongside Elliott Sober, he has advanced the thesis that the pertinent property of a group is that its members affect each others fitness, and that temporary aggregations of organisms have this property as well as isolated demes (Wilson, 1975a; Wilson and Sober, 1994). In cases where groups are aggregations of organisms that periodically mix, it has been shown that cooperation can indeed be promoted, despite the fact that selfish behaviour is advantageous *within* each group. That is, individual selection within groups favours selfishness, but this can be overcome by between-group selection favouring groups whose members are more cooperative, in the sense that those groups will contribute more individuals into the global population (Wilson, 1975a, 1987; Wilson and Colwell, 1981). One example of this is the “Haystack” aggregation and dispersal structure depicted in Figure 1.1. Indeed, Sober and Wilson argue that group selection, by this understanding, is both highly effective and pervasive in nature. This is because surely most organisms live in some kind of group in which fitness-affecting interactions take place (Krause and Ruxton, 2002), be that a family unit, or a habitat shared by unrelated conspecifics. Sober and Wilson therefore argue that group structure drives social evolution.

The response to this argument from many has been to agree that group structure affects social evolution, but not for the reasons that Sober and Wilson think. Rather, it is argued that such models are examples of *individual* selection in a group-structured population (Grafen, 1984; Nunney, 1985; Harvey et al., 1985; West et al., 2007a). That is, it pays individuals to be cooperative with other group members in certain situations. For example, if individuals with the same social behaviour tend to find themselves in the same groups, then it will pay for you to be a cooperator if that means you are on the receiving end of more cooperative behaviour from others (Sterelny, 1996; Maynard Smith, 1998). Since cooperation is individually advantageous in such scenarios, they see no role for group selection. Indeed, Wilson’s models can be redescribed in terms of inclusive fitness and kin selection, frameworks that were originally intended to avoid invoking group selection (Hamilton, 1963). On the other hand, Wilson and others stress that inclusive fitness and kin selection are instances of group selection, for they can only operate in group-structured populations, *sensu* Sober and Wilson (Hamilton, 1975; Wilson, 1975a; Wade, 1980; Michod, 1982).

How can this tension be resolved? Some authors have stressed that there are two different concepts of group selection at work in the literature (Damuth and Heisler, 1988; Mayo and Gilinsky, 1987; Arnold and Fristrup, 1982; Maynard Smith, 1987; Okasha, 2001). Type 1 group selection is invoked to explain the evolution of an individual-level trait, such as cooperative behaviour, in a group-structured population. That is, it explains how selection pressure at the group level causes an individual-level trait to increase in frequency in the population of *individuals*. By contrast, type 2 group selection is invoked to explain how a group-level trait increases in frequency in the metapopulation of *groups*, due to groups with the trait founding more offspring groups than those lacking it. Although the difference may at first seem subtle, it is a fundamental distinction (Okasha, 2001, 2006). In particular, only under group selection of type 2 do groups play a role in evolution strictly analogous to that of individuals (Arnold and Fristrup, 1982; Damuth and Heisler, 1988), and it has been argued that only in such cases can adaptations evolve that are genuinely "for the good of the group", in the sense that they evolved to increase fecundity and/or survival of the group *per se* (Maynard Smith, 1987). Everyone would agree that group selection of type 2 is genuine group selection, for it is natural selection at the individual-level frame shifted one level up the biological hierarchy, and surely very few would defend the position that natural selection can in principle only operate at the individual level (Lewontin, 1970). Moreover, a redescription of such a process in terms of individual selection will not be causally appropriate, since group fitness cannot be equated *directly* with individual fitness (Okasha, 2006). By contrast, in group selection of the type 1 variety, group fitness just *is* the fitness of the group members: this has to be the case, for type 1 selection explains why group-level selection pressures make one individual trait fitter than another, e.g., why cooperative rather than selfish behaviour evolves. Consequently, a redescription of any such models in terms of individual fitness must necessarily be available, for an individual trait can only be selectively advantageous if individuals with that trait are, on average, fitter than those without it (Wilson, 1975a; Dugatkin and Reeve, 1994; Sterelny, 1996; Sober and Wilson, 1998; Okasha, 2006). Despite Sober and Wilson's (1998) assertion that breaking down the selective forces on individuals into within- and between- group components provides the only correct causal account of evolution in such scenarios, few have been convinced of this (Nunney, 1985, 1998; Maynard Smith, 1987, 1998; Kitcher et al., 1990; Sterelny, 1996; Barrett and Godfrey-Smith, 2002; Kerr and Godfrey-Smith, 2002).

One might argue, then, that if everyone agreed to think only of type 2 group selection as "real" group selection then the problem would be solved. Yet, models of type 2 group selection are very rare in the literature, and we argue that this is because they presuppose that groups are evolutionary individuals in their own right. This then begs the question of how the groups came to be cohesive enough for this to be the case. On the other hand, type 1 group selection seems to be explained away by individual selection, for the groups can be treated as just a part of the selective environment. We will argue that social niche construction can bridge this impasse. In particular, if the groups in a type 1 scenario are being created by the evolving individuals themselves, then they cannot be treated as a static part of the environmental context. Type 1 group selection cannot, therefore, be seen as individuals merely adapting to their social environment, in the same way that they adapt to their abiotic environment in classical models of individual selection. We argue that what is interesting about type 1 is how the strength of it can increase over evolutionary time, as a result of individual adaptation. By strengthening the effect of group selection in this way, individuals create more cohesive groups, and hence start to move towards type 2 selection. Thus, we argue that treating groups in the same way as any other environmental property is unsatisfactory, given that group structure can itself change in an adaptive manner.

In the remainder of this chapter we elaborate on this argument, as follows. We first discuss group selection in its historical context, and how it was dismissed as an explanation for the evolution of cooperation in the 1960s and early 1970s. We then discuss the individualist response to how cooperation evolves through kin selection and increase of inclusive fitness (Hamilton, 1963, 1964a,b), before considering how they can be recast in terms of type 1 group selection, and the controversy that this has caused in the literature. We then turn to the distinction between type 1 and type 2 group selection, and how while it alone has failed to ease the tension between group- and individual-selectionists, the social niche construction framework can, alongside this, help to.

2.2 Early models of group selection and their failure

The idea that natural selection could potentially operate at a level higher than the individual was first postulated by Darwin (1871), in order to explain the evolution of human cooperative characters that benefited the whole tribe. The problem as Darwin saw it was that such characters might well confer no relative fitness advantage on their bearers compared to others in the same tribe, and so could not evolve by individual (within-tribe) selection. However, tribes with more cooperative individuals would surely outcompete (in terms of tribal warfare) those with less, and so cooperative traits could spread by between-tribe selection.

This school of thought continued through to the 1960s, whereby it was difficult to see how cooperative behaviours that benefit others, at an apparent cost to self, could evolve by individual selection, but easy to see how they might be favoured by group selection. Groups were seen as playing a role in evolution strictly analogous to that of individuals, with groups founding offspring groups and sometimes going extinct. A cooperative trait that benefited the group could then evolve because groups that had that trait would outcompete, in terms of founding more new groups and/or going extinct less frequently, groups that did not. Cooperation amongst individuals was therefore seen not as an adaptation of individuals, but of groups.

Attempts at explicitly modelling the spread of a cooperative trait by this mechanism were first made by Haldane (1932) and Wright (1945). However, neither were able to present a plausible model of the process. In particular, Wright envisaged a population composed of a number of reproductively isolated groups that occupied discrete geographical sites, but that were connected together by a flow of migrants, i.e., an island model. It was then supposed that a cooperative trait, which increased group productivity (in terms of number of migrants sent out) and longevity but which decreased the fitness of the bearer relative to others within the group, could reach fixation in a small number of groups by genetic drift. If such fixation of cooperation within one group occurred, then the possibility was laid open for cooperative behaviour to spread globally due to the increased productivity and longevity of that group. However, this process seemed inherently weak, since the rate of flow of migrants must be small enough so as not to disrupt the within-group process of genetic drift that might favour cooperation, but large enough so as to allow cooperative groups to export their progeny across the population (Wright 1945, rearticulated and placed into historical context in Wilson 1983b and Sober and Wilson 1998). Significantly, however, Wright only sketched out the above model; a quantitative or analytical analysis was not carried out.

Despite the lack of a believable model of group selection, and the exposition of the founders of neo-Darwinism that standard individual-level selection could not be assumed to lead to traits that increased group productivity (Fisher, 1930; Haldane, 1932), many ecologists and evolutionists employed the notion of group-level adaptations during the 1950s and 60s without thinking about a mechanistic process that could lead to them (Wilson 1983b and Okasha 2006 provide summaries of such thinking). Adaptations were postulated at the level of the group (Allee et al., 1949), population (Wynne-Edwards, 1962), species (Lorenz, 1963), or even the whole ecosystem (Dunbar, 1960), without considering how higher-level selection might be undermined by individual selection. This limitation of such thinking was exposed in the 1960s, however, when Williams (1966) launched an attack on group selection in general, and the thesis of Wynne-Edwards (1962) in particular.

Wynne-Edwards had proposed that populations evolved adaptations to avoid over exploiting their resources, for example, by reproductive restraint to control population density. Moreover, he proposed that since populations which over-exploited their resources would tend to go extinct, they would surely be replaced by other populations that had adaptations to prevent over exploitation. In this way, such population-beneficial adaptations could potentially spread by a process of selection between populations, just as individually-beneficial adaptations such as camouflage colouring spread by selection between individuals. Williams (1966) gave a highly influential argument, however, that such population-level selection was likely to be highly ineffective, since selection pressure at the population level for reproductive restraint would be opposed by selection pressure at the individual level for unabated reproduction. This follows from the fact that an unrestrained reproducer (arising by mutation or migration, for example) in a population of restrainers would leave more offspring. Therefore, assuming that unrestrained reproduction was a heritable trait, the frequency of unrestrained individuals would increase within the population with every generation; this is the problem of selfish cheating and “subversion from within” (Dawkins, 1976) that undermines group cooperation, and which was not considered by Wynne-Edwards. Furthermore, since the generation time of individuals is much shorter than that of populations, Williams claimed that competition between populations would hardly ever be able to counter this trend, since unrestrained mutants would be spreading quicker than populations were dying out, hence evolution would be faster at the individual than at the population level. More formally, natural selection at any level of the biological hierarchy requires a variance in fitness at that level. For example, under the popular Price Equation framework for studying group selection (Price, 1972; Hamilton, 1975), group selection trumps individual selection when the covariance between group character and group fitness is greater than the mean of the covariances between individual character and individual fitness within groups[1] (Okasha, 2005). In the kind of groups envisaged by Wynne-Edwards, i.e., whole populations, it is difficult to see how this condition can be met.

Arguments against the efficacy of higher-level selection became orthodox in evolutionary biology, and were supported by the first simulation models of demic group selection in the early 1970s[2] (e.g., Boorman and Levitt 1972; Levin and Kilmer 1974; see also Wade 1978 for a review). In these models, the groups were demes occupying resource patches. The productivity and/or survival rate of a deme was then proportional to the frequency of cooperative individuals that it contained. In this way, groups with fewer cooperators would more often be subject to extinction, and they would be replaced by migrants from groups with more cooperators. This was the kind of process that had been verbally described by Wynne-Edwards. The models, however, suggested that variance in fitness at the group level would only be sufficient to allow a cooperative trait (such as reproductive restraint) to evolve if the groups were very small. This is because it was assumed that cooperation would be driven extinct within all polymorphic groups, due to individual selection favouring selfishness (Maynard Smith, 1964, 1976), and the fact that the generation time of groups was much longer than that of individuals, giving individual selection much more time to act. If the groups were very small, however, then cooperation may be able to fix within some groups by genetic drift, as envisaged by Wright (1945). Cooperation could then potentially spread to other groups by migration. However, the intergroup migration rate typically had to be very low (on the order of two groups exchanging one migrant per generation (Maynard Smith, 1976)) in order to stop selfish cheats from “infecting” other groups too rapidly, which would further reduce between-group variance. The fact that the models only worked under such narrow conditions seemed to preclude group selection as a mechanism for the evolution of cooperative behaviour, except in extremely limited special cases.

---

[1] Where group character is defined as the mean character of the group members. So when considering cooperation, for example, individual character would be whether an individual is a cooperator or not, while the group character would be the proportion of cooperators. As it turns out, this relationship between group and individual character is fundamental when considering how to define group selection, as will be discussed later in this chapter.

[2] The dynamics of selection in a metapopulation where groups existed for multiple generations, exchanged migrants, and went extinct had resisted analytical treatment, so the availability of computer resources in the 1970s allowed a numerical treatment of the problem for the first time.

2.3 Kin selection and inclusive fitness as a mechanism for the evolution of cooperation

Despite the failure of group selection as a plausible evolutionary mechanism, it was still nevertheless the case that cooperative behaviour was to be found in nature and did appear adaptive, so the evolution of such traits did need to somehow be explained in terms of selective advantage. This was achieved in the 1960s by taking a gene's eye view of evolution, in which the fate of a gene, rather than the individuals that carry it, is followed (Hamilton, 1963, 1964a,b; Williams, 1966; Dawkins, 1976). A cooperative act that disadvantages the bearer, but that benefits others, will evolve if the net result of the action is to increase the frequency of genes coding for the cooperative act in the gene pool (Hamilton, 1963)[3]. This can occur if the benefits of cooperation fall upon other individuals that also carry the cooperative allele, for this can more than offset the fitness loss to then actor. This process is known as kin selection, a phrase coined by Maynard Smith (1964), for if social interactions take place between kin, then the interacting individuals will tend to share the same gene for cooperative behaviour, due to inheriting it from a common ancestor (Hamilton, 1963, 1964a; Maynard Smith, 1964).

The argument is made formal in Hamilton's (1964a) rule. This states that a social trait will evolve when $rb > c$, where b is the benefit of the cooperative act bestowed upon recipients, c is the fitness loss to the actor, and r is the genetic relatedness between actor and recipient (Hamilton, 1964a). Genetic relatedness was originally defined as the proportion of genes that actor and recipient of social behaviour share due to common ancestry, i.e., *genealogical* relatedness, but this concept has since been generalised, as discussed below. Hamilton's rule simply states that for a cooperative trait to spread, its effect on the recipient, discounted by relatedness, must be greater than the cost to the actor. Incorporating the effects on recipients, weighted by relatedness, in this manner yields the concept of inclusive fitness (Hamilton, 1964a,b), whereby both direct (the cost to the actor) and indirect (the benefits given to others) effects are used to calculate the fitness of an individual. Natural selection can then be viewed as acting to maximise individuals' inclusive fitness (e.g., Grafen 2006). Because inclusive fitness is a version of classical individual fitness that accounts for social interactions, kin selection is often seen as being fully compatible with individual selection (Maynard Smith, 1976). Alternatively, an equivalent direct fitness approach is available, where the benefits received from others are accounted for, rather than those given to others (Taylor and Frank, 1996; Frank, 1998); it is important not to double count and incorporate both, one must either account for the benefits given to others, discounted by relatedness, or the benefits received from others, where the relatedness gives the probability that they will be received (Grafen, 1984).

---

[3] We assume, for ease of exposition, that the trait is genetic. However, all that matters is that the trait is heritable, and this could occur if it was inherited culturally, rather than genetically, for example.

Under what kind of population structures is Hamilton's rule satisfied (and hence, in what kinds of population can cooperation evolve)? One immediately obvious candidate is family structured populations, where social interactions take place in kin groups, such as a bird's nest. Essentially, the offspring of an individual carrying a cooperative allele will also tend to carry that same allele, hence the relatedness between interacting individuals in a family group will be high. In particular, in diploid sexually reproducing organisms obeying Mendelian segregation, the relatedness between parent and offspring, or between siblings, is 0.5. This then gives the well-known result that an altruistic trait directed towards siblings or offspring can spread if the benefit is more than twice the cost, hence Haldane's quip that he would lay down his life to save two brothers (e.g., as cited by Sober and Wilson 1998). It was these kinds of population structure that Hamilton first envisaged, and which led Maynard Smith to relabel inclusive fitness as kin selection. Indeed, the relatedness term in Hamilton's rule was initially *only* interpreted in terms of genealogical ancestry, as defined above. Kin selection seemed much more plausible than traditional group selection as a mechanistic explanation for the evolution of cooperation, since many social interactions surely take place between relatives, thereby securing the necessary positive relatedness. By contrast, few populations seemed to exhibit the conditions for effective group selection as understood at the time kin selection theory was being developed, i.e., deme sizes of around 10 individuals and low migration rates (Maynard Smith, 1964, 1976; Levin and Kilmer, 1974). Kin selection therefore seemed pervasive in nature, while group selection appeared to be extremely unlikely to be effective in natural populations and hence to be of only limited interest as something that was theoretically possible (Wilson, 1983b; Sober and Wilson, 1998).

Moreover, later theoretical developments have shown that gene's eye thinking and inclusive fitness explanations for cooperation spread beyond kin selection, in the strict sense used above. In particular, what matters is not that the interacting individuals are kin, but that they are statistically likely to share the allele for cooperation (Hamilton, 1975; Grafen, 1985), i.e., that there is a positive regression on the frequency of the cooperative allele in actors and recipients (Queller, 1985; Frank, 1998). Consequently, relatedness need only be defined at the locus (or loci if cooperation is coded for by a suite of alleles) where the cooperative allele resides, rather than on a genome wide basis. In this way, inclusive fitness explanations for social behaviour are not limited to interactions with genealogical relatives.

What mechanisms, besides common ancestry, might produce the necessary positive regression between actor and recipient of cooperation required for an inclusive fitness explanation? Essentially, any mechanism that produces assortativity on behaviour will provide a positive regression. For example, individuals may be able to gauge the cooperative tendencies of others and use this information to choose with whom they interact. This may be achieved directly through cognition (Wilson and Dugatkin, 1997), or indirectly by means of a passive signal that cooperative individuals produce, and which is detected by other cooperators. In effect, such a signal would be the "green beard" gene of Hamilton (1964b) and Dawkins (1976), in which a cooperative allele has a pleiotropic effect that causes its bearers to emit a signal, and to only cooperate with other individuals also emitting that signal. Assortativity of behaviour without common ancestry could also arise if individuals with the cooperative allele somehow ended up living in the same habitats (Hamilton, 1975; Pepper, 2007). However, although all of these mechanisms can work in theory, kinship does seem theoretically to be the most robust way of securing positive relatedness (West et al., 2007a). This follows from the fact that if the interacting individuals are kin, then they share equal relatedness at all loci; thus, a cooperative allele that helps a genealogical relative benefits all other alleles on the same genome equally (Okasha, 2002; Leigh Jr., 2009). On the other hand, if the interacting individuals are not kin then alleles at loci unlinked to cooperation will not benefit, hence there could be selection pressure at all of these other loci for suppressing cooperation (Grafen, 1984; Okasha, 2002; Leigh Jr., 2009), in the same way that meiotic drive alleles can be suppressed by other alleles at unlinked loci (Alexander and Borgia, 1978; Leigh Jr., 2009).

2.4 The group selectionist response: new group selection models and kin selection as an instance of group selection

While the models of traditional demic group selection developed in the 1970s seemed very pessimistic for its role in social evolution, work was being undertaken at the same time by a number of theorists aimed at broadening our conception of group selection. In particular, D.S. Wilson (1975a; 1980) developed the notion of a trait-group – a subset of the population that have fitness affecting interactions with each other with respect to some trait, but not with organisms outside of that subset (with respect to the trait in question). Wilson, and others (Matessi and Jayakar, 1976; Uyenoyama and Feldman, 1980), have argued that such trait groups do not even have to be reproductively isolated, but instead can be just fleeting aggregations of organisms that interact during some part of their life-cycle, such as birds that breed communally (Brown, 1978), or bark beetles that aggregate to attack a tree (Avilés et al., 2002). On the other hand, trait-groups can also last for many generations, for example, they could consist of organisms that share a common resource patch and reproduce on that patch for a number of generations, until the resource becomes depleted (Wilson and Colwell, 1981; Wilson, 1987). Thus, trait-groups are intended by Sober & Wilson as a generalisation of the evolutionary concept of group from the demes considered in early group selection models, for they see the relation between individuals of "interacts with" as the pertinent one for defining groups that have evolutionary consequences (Wilson and Sober, 1994; Sober and Wilson, 1998).

How can the presence of trait-group structure in a population affect the evolution of social traits? If organisms disperse from their trait-group at some stage and mix freely back into the global population then cooperative behaviours can be selected for. This is because trait-groups with a greater proportion of a cooperative allele will contribute more individuals into the population than those with a greater proportion of selfish individuals. Crucially, this can occur even if cooperation is selected *against* within each trait-group, since the difference in the number of individuals produced by groups at the dispersal stage can more than offset the decline in cooperation within each group. By this mechanism, cooperation can decline in frequency *within* each and every trait-group but still increase in frequency in the *global* population, a result that is an instance of the statistical phenomenon known as Simpson's Paradox (Simpson, 1951; Sober and Wilson, 1998). An example of this process was provided in Figure 1.1, in which organisms reproduced on isolated patches for a number of generations, before a global mixing stage formed the next generation of groups.

Wilson (1975a, 1980), Sober (1984, 1987), and others maintain that this process clearly demonstrates group selection. That is, they argue that two selective processes act in trait-group structured populations. They claim that *individual* selection favours selfishness; this is the reason that the selfish type increases in frequency within each trait-group. That is, they equate individual selection with within-group selection. Following on from this, they claim that it is *group* selection that allows cooperation to evolve in such populations, since cooperation only evolves because trait-groups with a greater proportion of cooperators contribute more individuals into the global population. Thus, just as individual selection is the differential productivity of individuals, so group selection is the differential productivity of trait-groups, they argue (Wilson and Sober, 1989, 1994). Moreover, they argue that natural selection at any level requires variation between the entities at that level (Wilson, 1997), and that this is clearly the case with trait-group selection. In particular, it has been proved that the greater the variance in initial trait-group composition, the more effective trait-group selection is at promoting cooperation (Price, 1972; Wilson, 1975a; Hamilton, 1975). In the limit, if all trait-groups have the same frequency of a cooperative allele, then cooperation cannot evolve, for there would not be any difference in group productivity at the dispersal stage, and hence no component of selection between trait-groups (Wilson, 1990, 2004).

More formally, Wilson (1975a, 1977, 1979, 1980) has shown that if trait-groups stay together for a single organismic generation before dispersal, then random variation between trait-groups is sufficient to allow a class of cooperative behaviours known as "weak altruism" (Wilson, 1979, 1980) to evolve. These are behaviours that increase the number of offspring of the actor, but increase the number of offspring of other group members by even more. Examples of these traits can include both the public goods production and consumption behaviours discussed in Chapter 1. Although these traits increase the number of offspring of the actor, Wilson (1979, 1990) stresses that they nevertheless require group selection to evolve, for they put the actor at a local disadvantage compared to other group members, and this local disadvantage can only be compensated for by differential group productivity[4]. In particular, these traits cannot evolve if between-group variance is less than binomial (Wilson, 1990, 2004). If between-group variance is greater than random, then so-called "strong altruism" can evolve; these are traits that decrease the number of offspring of the actor, while increasing those of other group members. These are of the kinds of altruistic traits considered by Hamilton (1964a,b), and include such extreme forms of cooperation as the suicidal bee sting, or worker sterility in ant colonies. Wilson (1977, 1980) suggests that random variation in group composition is likely to be a lower threshold in most natural populations. Mechanisms such as assortative interactions between individuals of the same type (Wilson and Dugatkin, 1997) can produce above random between-group variance. Likewise, special mechanisms would be required to produce below random between-group variance, for this would require a mechanism of disassortativity (Wilson, 1977). However, in cases where the groups stay together for multiple organismic generations before dispersal, then the initial variance when the groups are formed can be either magnified or reduced by the action of individual reproduction and selection within the groups (Wilson and Colwell, 1981; Wilson, 1987; Fletcher and Zwick, 2004, 2007). Consequently, random variance in initial group composition is no longer the threshold on what type of cooperation can evolve in multi-generational groups (Fletcher and Zwick, 2004), and the hard distinction between weak and strong altruism is no longer relevant (Fletcher and Doebeli, 2006; Fletcher and Zwick, 2007).

The dispersal stage in trait-group models is crucial to maintaining the initial between-group variance, for selection within each group will decrease the frequency of the cooperative trait towards extinction. The groups must therefore mix and be reformed before cooperation has declined in frequency by too much within each group, so that differential group productivity at the mixing stage can more than offset this decline. It is this regular group reformation that prevents selfish types from fixing in groups, and hence the problem of cooperation being "subverted from within" described by Williams (1966) and Dawkins (1976), and which hampered the evolution of cooperation in the traditional demic models (e.g., Maynard Smith 1964; Levin and Kilmer 1974). For example, in his original "Haystack" model (analogous to that shown in Figure 1.1), Maynard Smith (1964) assumed that cooperation must be driven extinct in any group that contained the selfish allele. Wilson (1987) argued that dispersal can take place before this occurs, for surely it will take many generations for a cooperative allele to be driven extinct. Wilson showed that when dispersal occurs more frequently, then group selection is a much more effective force. More generally, group selection is effective when the "generation time" of groups is relatively short, for then between-group selection will take

---

[4] This argument of weak altruism requiring group selection to evolve has been contested in the literature (Nunney, 1985, 1998; Maynard Smith, 1998), as we discuss later in this chapter.

place more frequently. The generation time of the populations of red grouse described by Wynne-Edwards (1962) would presumably be very long compared to that of the individuals within a population, hence group selection would be very weak compared to individual selection in that case (Maynard Smith, 1964; Williams, 1966); this was shown more formally by the traditional demic models of group selection (Boorman and Levitt, 1972; Levin and Kilmer, 1974). Wilson and Sober's point, however, is that such groups are extremes, and that group selection is much more likely in other kinds of group.

Strikingly, Wilson (1977) and others (Wade, 1980; Michod, 1982; Breden, 1990) assert that one such effective trait-group structure is the sibling groups of classic kin selection models. That is, they see kin selection and inclusive fitness explanations for the evolution of cooperation as *instances* of group selection. How is this so? The argument turns on the fact that an inclusive fitness or kin selection explanation for the evolution of cooperation itself requires localised interactions, for genetic relatedness is the difference between the frequency of a trait experienced within interaction groups, compared to the frequency in the global population (Grafen, 1985). Indeed, relatedness is formally equivalent to the proportion of the total genetic variance in a trait that is *between* groups, i.e., it is the between-group variance (Queller, 1992b). Moreover, the net result of selection in this interaction-group structure can naturally be broken down into two components by the Price equation (Price, 1972; Hamilton, 1975; Breden, 1990). Specifically, in kin selection models a within kin-group component favours selfishness, for selfish individuals receive the same benefits as other members of their kin group without paying the cost; that the group members are kin does not change this basic fact that follows from the definitions of cooperative and selfish behaviour. Thus, kinship alone does not remove the problem of selfish cheating, and so the differential productivity of *individuals* within a kin group selects for selfish behaviour (Wade, 1980; Michod, 1982). Conversely, a between kin-group component favours cooperation through the differential productivity of *kin groups*: kin groups with a greater proportion of cooperators contribute more individuals into the population, in exactly the same way as other trait-group models. Thus, kin selection is itself an instance of Simpson's Paradox, for cooperation declines in frequency locally but can nevertheless spread globally, exactly as in any other trait-group selection model. In general kin selection is, by this argument, highly effective at promoting cooperation because a kin-group structure automatically provides a high between-group variance (Wilson, 1977, 1980). In particular, a single kin group will tend towards genetic homogeneity due to common ancestry, thereby reducing within-group variance and increasing between-group variance (Wilson, 1975a, 1977). Just like any other form of group selection, the efficacy of kin selection is increased by factors that reduce the within kin-group variance and hence increase the between kin-group variance; such factors include inbreeding (Wade and Breden, 1981; Breden and Wade, 1981), and monogamous relationships between parents (Peck and Feldman, 1988).

Not only can kin selection be understood in terms of group selection, but this understanding has itself furthered kin selection theory (Queller, 1992a, 2004). In particular, it was originally thought that a viscous population, where siblings remain in their place of birth throughout their lives, would be particularly conducive to kin selection, by ensuring a high relatedness between interacting individuals (Hamilton, 1964b; West et al., 2002). However, it has since been proved that unless cooperative interactions take place on a smaller scale than those for resource competition, cooperation cannot evolve despite high relatedness (see Taylor 1992 for the analytical proof based on Hamilton's rule, and Pollock 1983, Goodnight 1992, Wilson et al. 1992 and Mitteldorf and Wilson 2000 for simulation models; see also di Paolo 2000). This is because if the interacting individuals also permanently compete for the same resource, then this competition can remove selection pressure for helping those individuals, even if relatedness is very high (Pollock, 1983; Taylor, 1992; West et al., 2002). That is, within-group selection favours selfishness locally (Pollock, 1983; Goodnight, 1992; Wilson et al., 1992), and within-group selection here manifests itself through local resource competition. Cooperation can consequently only spread by differential group productivity, which cannot be realised in a purely viscous population, because viscosity imposes local population density regulation that prevents more cooperative groups from growing to a larger size than more selfish groups [5] (Wilson et al., 1992). This is exactly the same as the argument given above for why groups must be periodically reformed in order for group selection to be effective: if more cooperative groups cannot realise their increased productivity through some form of dispersal, then cooperation will continue to decline in frequency within each group and, consequently, decline globally (Wilson et al., 1992).

2.5 Individualist attacks on the new group selection models

2.5.1 Group selection as individual selection in a group context

When trait-group selection theory was first developed, its compatibility with individual selection and neo-Darwinism was stressed (Wilson, 1975a, 1980). This follows from the fact that the fittest groups contribute the most *individuals* into the population, hence group fitness is the mean (or sum of the) fitness of the individual group members (Mayo and Gilinsky, 1987; Damuth and Heisler, 1988; Okasha, 2001). Because of this, when cooperation evolves it does so because cooperative individuals are fitter than selfish individuals when averaged across all groups, i.e., averaged across the whole population (Wilson and Sober, 1994; Sober and Wilson, 1998). Sober and Wilson (1998) argue, however, that this in no way undermines the role of group selection, for if cooperation

---

[5] Think, for example, of a computer simulation on a lattice, where all of the sites on the lattice become occupied.

is to evolve then it of course must increase in frequency in the global population of individuals. They contend that to average individual fitness over all groups, and then to say that cooperation is on average fitter, is fallacious, for it ignores the fact that cooperation is locally disadvantageous and spreads through differential group productivity. In particular, it does not matter if cooperation is fitter at some point in time when averaged over all groups unless there is some form of dispersal, for otherwise this fitness advantage of cooperation will only be transitory (Sober and Wilson, 1998), as was discussed above in the case of a viscous population.

Others, however, have argued that an individual selection description of evolution in group-structured populations is just as valid, and is not fallacious (Maynard Smith, 1987, 1998; Kitcher et al., 1990; Sterelny, 1996; Barrett and Godfrey-Smith, 2002; Kerr and Godfrey-Smith, 2002; Gildenhuys, 2003; Foster et al., 2006a; West et al., 2008). In particular, it is argued that such "group selection" models can be viewed as a case of ordinary frequency-dependent individual selection, since being a cooperator is individually advantageous when the population structure provides a sufficient assortment of behaviours, i.e., when there is a sufficient clustering of cooperators into groups (Sterelny, 1996; Maynard Smith, 1998; Skyrms, 2002). The groups can then be viewed as merely providing the *context* in which this frequency-dependent selection takes place, rather than as units of selection in their own right (Dugatkin and Reeve, 1994; Sterelny, 1996; Kerr and Godfrey-Smith, 2002). In contrast to the (trait) group selection approach, which sees both within- and between-group processes as drivers of gene frequency change, this argument sees only the global dynamic of frequency-dependent selection as important. Local dynamics within groups are not explicitly accounted for, for example, inclusive fitness equations do not contain an explicit notion of group fitness. Such arguments find favour with proponents of a reductionist philosophy of science (Okasha, 2005, 2006), including Dawkins (1976) and Maynard Smith (1998), for they raise the question as to why any group-level explanations of social evolution are needed.

One particular instance of the "individual selection doing away with group selection" argument concerns the evolution of weak altruism (Wilson, 1979, 1990) in randomly formed groups that stay together for only a single generation before dispersal. Recall that weakly altruistic traits increase the number of offspring of the actor, but increase those of other individuals in the same group by even more. Because they increase the offspring count of others by even more, Wilson (1979, 1980, 1990) argues that they cannot evolve by individual selection, and so must evolve by group selection. That is, they evolve because of their beneficial effect on other group members. Nunney (1985), however, argues that this is not the case, for if groups are reformed *randomly* every generation, then the effect of a trait on other group members cannot matter. This is because each behavioural type will experience each possible group composition with the same frequency, i.e., when averaged over all groups, each type will experience the same behavioural frequencies as exist in the global population. In effect, individuals are placed in new random group contexts every generation, and so the context cannot influence selection, Nunney argues. As a result, the direct effect of the trait on self is left as the only pertinent evolutionary force, and so in single-generational trait-groups any trait that increases the number of offspring of the bearer is selected for, regardless of its effect on others (Nunney, 1985; Wilson, 2004). Thus Nunney, alongside others (Foster et al., 2006a; West et al., 2007a), asserts that weak altruism evolves by *individual* selection. Significantly, however, this argument does not hold when randomly formed groups stay together for more than a single generation, because then each type will not, on average, experience the same behavioural frequencies as in the global population (Fletcher and Zwick, 2004). Rather, although the group context is initially random, after the first generation (descendants of) the individuals carry on interacting in the same context, and so this context *does* matter.

2.5.2 The gene as the fundamental unit of selection

There is a tendency in some circles to dismiss group selection by claiming that the gene is the fundamental unit of selection (Williams, 1966; Dawkins, 1976). The argument goes that only genes copy themselves with sufficient fidelity to cause sustained evolutionary change in a population, for entities further up the biological hierarchy, such as sexually reproducing organisms, do not pass copies of themselves intact into the next generation (Williams, 1966; Dawkins, 1976). If entities are not passed on intact through the generations, then sustained selection cannot, over more than one generation, favour one entity over another, for the targets of selection will be changing with every generation (Williams, 1966; Dawkins, 1976). Thus, only units with high copying fidelity, such as genes, can be the true targets of selection, it is argued (Williams, 1966; Dawkins, 1976). Consequently, all selection must ultimately be “for the good of” genes, and not entities further up the biological hierarchy such as individual organisms or groups (Williams, 1966; Dawkins, 1976).

Linked with this is the idea that all entities further up the biological hierarchy are made of genes, so any adaptation that increases the frequency of entities at level x must necessarily increase the frequency of the genes that created those entities. Essentially, this is a restatement of the standard definition of evolution as change in gene frequency (Fisher 1930, see Sober 1984 and Okasha 2006 for discussions of this point). However, this fact cannot be used as an argument against the existence of group selection without also arguing against selection at *any* other level of the biological hierarchy, including the individual level (Sober and Wilson, 1998). That is, both individual and group selection will cause some *genes* to increase in frequency at the expense of others. Moreover, genes are not directly subject to natural selection, but are only indirectly selected through the success of the phenotypes that they create (Mayr, 1984; Sober and Lewontin, 1982; Sober, 1984). These phenotypes can be manifested in individuals or groups (Sober and Wilson, 1998). Thus, there is a fundamental distinction between the *replicators* that copy themselves with high fidelity and so pass on their structure intact through the generations (genes), and the *vehicles* whose phenotypes are directly exposed to selection[6] (individuals, groups, or any other level of the biological hierarchy) (Dawkins, 1976, 1984; Hull, 1984). Consequently, arguing that only genes are replicators does not in any way undermine the role of either individuals or groups as vehicles of selection, and so "Selfish Gene Theory" (Dawkins, 1976) cannot be used to rule out group selection *a priori*, as is sometimes thought (Sober and Wilson, 1998). On the other hand the gene's eye view to evolution, whereby the fate of single alleles in the population is tracked, can provide another perspective in which to view group selection, for group selection will itself result in a change in gene frequencies[7] (a change in gene frequencies due to differential group productivity and/or extinction is how Wade 1978, and others, *define* group selection). Thus, the results of group selection can themselves be measured in the currency of gene frequencies, exactly as the results of individual selection can be. Indeed, we do so throughout this dissertation.

The notion that the gene is the fundamental unit of selection, then, is not an empirical thesis, but rather one about how we should best represent the selective process (Sober, 1990). This should be contrasted with another meaning of genic selection that is present in the literature. This is the idea that genes within individuals can compete with each other, just like individuals can compete with each other within groups (Wilson and Sober, 1989; Wilson, 1997). Thus, selection on genes need not necessarily be in alignment with selection on individuals, just like selection on individuals need not necessarily be in alignment with selection on groups. Genic selection not being aligned with individual selection is demonstrated by the existence of meiotic drive genes, that increase their representation in the gamete pool at the cost of a reduction in fitness of their individual carrier (Alexander and Borgia, 1978; Leigh Jr., 2009). Thus, meiotic drive genes increase their own genic fitness, but lower the individual fitness of their carrier. The strength of such genic selection is then an empirical matter that must be determined on a case-by-case basis, and this is the way that gene-level selection is understood in the multi-level selection literature (Wilson, 1997; Sober and Wilson, 1998; Okasha, 2006). That is, genes are one level of the biological hierarchy at which selection can act, alongside individuals and groups (Wilson, 1997). This is in contrast to the representational thesis of genic selection, which simply states that where genes are the particles of heredity, selection at any level can be described in terms of gene frequency change (Sterelny and Kitcher, 1988).

The Selfish Gene Theory of Dawkins is the representational thesis of genic selection (Sober, 1990). Within this framework, both groups and individuals play the role of

---

[6] Dawkins uses the terms replicator and vehicle, Hull uses interactor in place of vehicle.

[7] We assume that the replicators are genetic throughout this section for ease of exposition. However, there are other kinds of replicators, for example, cultural memes (Dawkins, 1976; Richerson and Boyd, 2005).

vehicles of selection. There is, however, a caveat. Dawkins (1982, 1994), has resisted the argument that many kinds of groups are vehicles, stating that groups typically lack the necessary cohesion, i.e., they do not present themselves as a single unit to selection in the way that individuals do. Wilson and Sober (1989, 1994) argue that such cohesion must itself be an evolved property, and that more cohesive groups must be the product of a process of group selection, for more cohesive groups presumably outcompete less cohesive ones.

Sterelny (1996) has argued that the most convincing examples of groups as vehicles are those where group fitness is not just the simple sum of individual fitnesses. He argues that in such cases, it is possible to distinguish between adaptations that evolved for the good for the group and those that evolved for the good of the individual. To support this, he gives an example of a sister-killing gene that has evolved in some eusocial insect species, which induces a queen to kill any other queens that are born in the same colony. Sterelny proposes two possible hypotheses for how this gene could have evolved. The first is that it evolved because it reduced genetic variation within the colony. This would reduce within-colony conflict as the colony developed, since all colony members would be descendants of the single queen, thereby reducing the potential for within-colony selection. Such reduced within-colony selection would help that colony in competition with others, by favouring more cooperative interactions between the colony members. The second hypothesis is that once the sister-killing gene arose by mutation, it would rapidly be selected for by individual selection. This is because any queen that did not kill others would loose out and fail to pass its genes on, i.e., sister-killing could be seen as defect in a Prisoner's Dilemma game. In this case, the gene would spread by individual selection, even though it might be detrimental to colony output. Both of these theses describe alternative mechanisms by which a gene for sister-killing could increase its representation in the gene pool. Sterelny argues that the first hypothesis would result in a group-level adaptation, the second in an individual-level adaptation. Moreover these two hypotheses, and hence who the adaptation benefits, could be discriminated between empirically. Sterelny concludes that if sister-killing did indeed evolve because it reduced within-colony variance, then it is much more convincing to view the colony as a vehicle of selection in its own right.

By Sterelny's argument, then, trait-groups would not count as vehicles of selection because trait-group fitness just *is* the fitness of the group members. Individuals therefore seem to be the vehicles, not the trait-groups. This links with the argument about "groups as the context for individual selection" discussed in Section 2.5.1, for if individuals rather than groups are the vehicles in a group-structured population, then surely an individual selection story is sufficient, and group fitness does not need to be invoked (Sterelny, 1996). This argument applies to the construal of kin selection as an instance of trait-group selection discussed in Section 2.4. That is, if kin groups are not vehicles of selection, then why do we need to see selection as acting on the kin group as a whole? For example, Maynard Smith (1987, pp.123-124) wrote:

> When Wilson (1975) introduced his trait-group model, I was for a long time bewildered by his wish to treat it as a case of group selection, and doubly so by the fact that his original model...had interesting results only when the members of groups were genetically related, a process I had been calling kin selection for ten years.

In other words, Maynard Smith thought that trait-group selection was just a case of individual selection (see also Maynard Smith 1998), and was very different from the original formulation discussed by Wynne-Edwards (1962), which had sparked the group selection controversy. Sober & Wilson, on the other hand, see selection on trait-groups and selection on the breeding populations of red grouse described by Wynne-Edwards as part of a continuum, and they argue that both types can be vehicles of selection and can evolve adaptations "for their own good" (Wilson and Sober, 1994; Sober and Wilson, 1998). By contrast, Maynard Smith (1987) argued that trait-groups cannot evolve adaptations, because they lack heredity and so are not "units of evolution", that is, units that have the properties of multiplication, variation, and heredity, and are hence subject to natural selection (Maynard Smith, 1987). In a similar vein, authors such as West et al. (2007a, 2008) see a clear distinction between the early demic models of group selection that were ineffective, and the newer trait-group models. They argue that only the former represent real group selection, and that consequently, real group selection is just as unlikely to occur as was originally thought when first discussing Wynne-Edwards' hypothesis. They then view trait-group selection as an instance of kin selection, rather than the other way around. In the next section, we discuss suggestions for how this tension can be resolved that have been put forward in the literature, and why in general they have not succeeded. We then turn to how social niche construction can build a bridge between individual- and group- selectionists.

2.6 Resolving the tension between individual and group selectionists

Let us first consider again the relationship between kin and group selection. In Section 2.4, we detailed an argument that kin selection is actually a particular instance of group selection, i.e., that kin selection is really selection at the level of kin *groups* (Wilson, 1975a, 1977; Wade, 1980; Michod, 1982). However, as discussed above, some authors prefer to see this the other way around, with kin selection being the fundamental process of which trait-group selection is an instance (Foster et al., 2006a,b; West et al., 2007a, 2008). Thus, a group selectionist would say that kin selection only works through the differential productivity of kin groups, which requires between-group variance and so is clearly a group selection process. On the other hand, an individual selectionist would say that kin selection works on inclusive individual fitness, and that cooperation only evolves in group selection models when there is a positive coefficient of relatedness between the group members (Maynard Smith, 1987). We now turn to how this debate might be resolved.

2.6.1 Pluralism between group and kin selection

A consensus has emerged amongst some authors that there is no one approach that is fundamentally correct, for both have heuristic value, they argue (Dugatkin and Reeve, 1994; Kerr and Godfrey-Smith, 2002). The argument is that both kin and group selection are equally valid descriptions of how cooperation evolves, and each can provide different insights (Kerr and Godfrey-Smith, 2002). For example, kin selection models have proved particularly useful for studying conflicts in the social insects (Queller, 2004; Boomsma and Franks, 2006), and the concept of genetic relatedness can be more intuitive than between-group variance to many field ecologists (Grafen, 1984), and is easier to measure in the lab thanks to modern genetic marker techniques (Queller and Goodnight, 1989) (West et al. 2007a discusses these and further heuristic advantages of a kin selection approach). On the other hand, group selection thinking has highlighted the local disadvantage of altruism even amongst immediate genealogical relatives, which is why purely viscous populations are not conducive to the evolution of cooperation (Wilson et al., 1992; Queller, 2004). It has also stimulated the creation of models where *genealogical* relatedness is not the driver of strong altruism (e.g., Peck 2004; Santos and Szathmáry 2008).

This pluralistic position may seem attractive, but it has not settled the group selection debate, for three reasons. The first reason is that authors on both sides still advocate that their way of looking at social evolution is the correct one (Sober and Wilson, 1998; West et al., 2007a, 2008; Wilson, 2007). That is, they are able to acknowledge the other position without giving it parity. The number of *real* pluralists in the social evolution field is thus very small. The second reason is that the formal theory of individualists rests on the calculation of inclusive fitness in order to determine the direction of selection on a social trait. This theory, however, is based on pairwise interactions between individuals (van Veelen, 2009). van Veelen (2009) has shown that if interactions instead involve more individuals then inclusive fitness does not correctly predict the direction of selection, unless the interaction takes the form of a linear public goods game[8] . Such linearity

---

[8] The crucial point is that more than two individuals interact simultaneously. In many models, although the trait-group size is larger than 2, the interactions are still assumed to be pairwise. For example, in the models presented in this dissertation, 2 player games are generalised to trait-groups of size n, by multiplying the 2x2 payoff-matrix by the proportion of cooperative and selfish individuals in the group, respectively. The interpretation of this is that pairs of individuals in the group interact randomly. Alternatively, an n-player game could be formulated as an nxn payoff matrix, representing all

in the costs and benefits of cooperating is a special case, and only applies when the benefit of cooperating does not change with the frequency of cooperators in the group. van Veelen provides many examples of models in the literature with non-linear costs and benefits, and argues that such non-linearity is likely to be relevant during many of the major transitions to new levels of biological organisation. This suggests that inclusive fitness may be the wrong methodology for modelling the major transitions, and that multi-level selection methods may instead be necessary to correctly analyse models, since they do not suffer from problems with non-linearity when the trait-group size is larger than 2.

The third problem with a pluralistic approach is that it does not touch on the question of whether trait- and kin- group selection are real group selection or not. That is, are these selective processes the same kind that Wynne-Edwards envisaged, and which the demic group selection models attempted to capture? This question is important, because there is a strong tendency in the literature to distinguish between "old" and "new" group selection. "New group selection" is considered to be the trait-group style models, whereas "old group selection" is that envisaged by Wynne-Edwards, according to some authors (Grafen, 1984; Reeve and Keller, 1999; West et al., 2007a, 2008). Old group selection is thought by these authors to be fundamentally different to individual selection, and is thought to have been ruled out as a plausible evolutionary mechanism because, as shown by the demic models of the 1970s (Boorman and Levitt, 1972; Levin and Kilmer, 1974), it was ineffective under plausible conditions. These authors then redescribe trait-group selection in individualist terms, and hence see no real role for group selection of any kind (West et al., 2007a, 2008).

2.6.2 Old versus new, and type 1 versus type 2, group selection

It is important to carefully evaluate the claim that old and new group selection are fundamentally different selective processes, for much of the modern group selection debate turns on this issue. In the philosophical literature, a distinction is made between type 1 and type 2 group selection models (Damuth and Heisler, 1988). Type 1 group selection models aim to explain the evolution of individual traits in a group-structured population (Damuth and Heisler, 1988; Okasha, 2001, 2006). Groups are assigned fitnesses in such explanations, but fitness is measured in terms of the number of *individuals* that the group contributes into the global population (Damuth and Heisler, 1988; Mayo and Gilinsky, 1987; Okasha, 2006). This has to be the case, given that such models aim to explain how an individual trait, such as cooperation, increases in frequency in the population of *individuals* (Okasha, 2001, 2006). So individuals, and not groups *per se*, are the

---

individuals in the group participating in one simultaneous interaction. It is in these cases that inclusive fitness does not give the right answer if the group size is greater than 2, and the interaction is not a linear public goods game.

focal units in such models (Damuth and Heisler, 1988), and this is why a redescription of such models in terms of individual fitness and selection is always available.

By contrast, type 2 group selection models explain the evolution of group-level traits (Damuth and Heisler, 1988; Okasha, 2006). Consequently, group fitness has to be measured in terms of the number of offspring *groups* founded by a parent group (Damuth and Heisler, 1988; Mayo and Gilinsky, 1987; Okasha, 2006), for we wish to explain how a group-level trait increases in frequency in the metapopulation of groups. In such models we therefore need to track the frequency of groups as well as individuals, so the groups *are* focal units in this variety of group selection. It is important to stress that it is not a matter of perspective as to whether we count groups or individuals or both (Okasha, 2006): if we are interested in the evolution of an individual-level trait, then we need only count individuals, but if we are interested in the spread of a group-level trait, then we need to track the frequency of groups with that trait. In type 2 group selection we will often need to track the frequency of individuals as well, because selection on individuals within a group may be in opposition to selection between groups that promotes the group-level trait. This does not, however, undermine the need to track groups also.

An example of this is provided in the Stochastic Corrector model of Szathmáry and Demeter (1987). In this model, different types of replicator are encapsulated inside protocells. Within a protocell, the replicators contribute to a common metabolism, but use that metabolism at different rates. Thus, selection on replicators within protocells will favour the type that uses the shared metabolism at the fastest rate, ultimately tending to competitive exclusion of all other replicator types within the protocell (and in fact, the death of the protocell, as protocell functionality is assumed to require the presence of all replicator types (Grey et al., 1995); this is thus a perfect instance of the Tragedy of the Commons (Hardin, 1968)). But, there is an optimum combination of different replicator types, which causes protocells close to that combination to divide more rapidly, and hence to become the more common protocell type[9]. The model thus explains how the protocell-level trait of having a certain combination of replicators increases in frequency in the population of protocells. In order to do so, the model needs to track both the frequencies of different types of replicator within protocells, and the frequencies of protocells that have different combinations of replicators. The model is therefore of the type 2 group selection variety.

It seems clear that all authors in the literature would agree that type 2 selection is “real” group selection, for it is an exact analogue of individual selection shifted one level up

---

[9] The model is called the “Stochastic Corrector” because there are two forms of stochasticity that generate the necessary between-protocell variance in replicator composition, for group selection to act on. These are the random assortment of replicators into offspring protocells when the parent divides, and a stochastic component of reproduction within protocells owing to the small number of replicators present (Szathmáry and Demeter, 1987; Szathmáry, 1999). The between-group variance generated by this stochasticity allows group selection to “correct” for the tendency of within-group selection to lead to competitive exclusion, and hence the loss of replicator types required for protocell functionality (Szathmáry, 1999).

the biological hierarchy. Moreover, this is the kind that authors such as Grafen (1984) and West et al. (2007a, 2008) equate with "old" Wynne-Edwards style group selection. However, was type 2 group selection really what Wynne-Edwards was postulating? Several authors attempted to model the Wynne-Edwards process (Maynard Smith, 1964; Boorman and Levitt, 1972; Levin and Kilmer, 1974), by considering that demes with a greater proportion of altruists would be less likely to go extinct and/or would send out more individuals to colonise vacant resource patches. On close inspection, these models are actually of type 1 (Okasha, 2001, 2006). In retrospect, this had to be the case, for they modelled how the individual-level trait of altruism could spread as a result of differential group productivity and/or extinction (Okasha, 2001). That is, they showed conditions under which an altruistic allele could increase in frequency in the population of *individuals*, due to differential group productivity and/or extinction (Maynard Smith, 1964; Boorman and Levitt, 1972; Levin and Kilmer, 1974). None of these models tracked the frequency of particular kinds of *group* (i.e., the frequency of groups with $x\%$ of altruists in the metapopulation of groups). Their results were framed entirely in terms of the frequency of alleles carried by individual organisms, for this is all that is needed to explain the evolution of an individual phenotype like altruistic behaviour. This should be contrasted with the Stochastic Corrector model, which explains the evolution of a group phenotype[10], namely, the combination of replicator types inside the protocell.

Okasha (2001) points out that the demic group selection models, which aimed to model the Wynne-Edwards process, could instead have modelled the evolution of a group-level phenotype "group has $x\%$ altruists", rather than the individual-level phenotype "individual is an altruist". They did not do so, however, for their focus was on how individual altruists could increase in frequency in the population. In general, a group-level trait must depend on individual-level traits, which in turn must depend on the genes that code for them. But, this does not mean that we do not need to explain the evolution of a trait at a higher level. For example, the Stochastic Corrector model was designed to explain how protcells with the optimum (from the point-of-view of the protocell) combination of replicator types came to increase in frequency in the population of protocells. Although this trait depends on the replicator-level trait "is a replicator of type x", the aim of the model was *not* to explain why a certain proportion of replicators, in the global population of replicators, evolved to be of type x. The fact that changes in protocell frequency result in global changes in replicator frequency does not preclude them from being different questions. On the other hand, the demic group selection models cited above did just seek to explain the evolution of an individual-trait, and the

---

[10] The use of group- and individual- phenotype, rather than trait, emphasises the fact that in either case, it is the product of genetic interactions with the biotic and abiotic environment. Thus, group phenotypes have a genetic basis, just like individual phenotypes. The fact that groups are made of individuals does not prevent groups from having their own phenotypes, in the same way that the fact that individuals are made of cells does not prevent individuals from having their own phenotypes. This follows on from the separation between replicators and vehicles made by Dawkins (1984) and Hull (1984), for they argued that the fact that entities at level x of the biological hierarchy are made of genes does not prevent those entities from having phenotypes and undergoing selection.

fact that this resulted in the composition of groups changing was a by-product. This is why "an individual selection in the context of groups" account of such evolution is available.

So, the early group selection models were indeed of type 1 (Okasha, 2001, 2006). That is, they are of the same type as trait-group models and kin selection models. They can therefore be viewed as "individual selection in a group context" in the same way that trait-group models can be, contrary to what Grafen (1984) and West et al. (2007a, 2008) suggest. These demic models were ineffective not because they were "old" group selection, but because they lacked effective mechanisms to generate between-group variance (such as kinship), and preserve it (such as global dispersal stage). On the other hand, the"old" style group selection Stochastic Corrector model is effective because the fact that protocells regularly split allows between-group variance to be regenerated (Szathmáry, 1999). The connection between the early group selection models and those of the trait-group variety is particularly vivid when one considers multi-generational trait group models, such as those in Wilson (1987) and Fletcher and Zwick (2004). Arguments that trait-group models are not "real" group selection, but merely instances of individual selection, thus apply to any of these models, and indeed must do so since all of them consider the evolution of individual traits.

What, then, of type 2 models? Even individualists such as Grafen & West would presumably agree that an individual-level selection explanation is much less appropriate in such cases. Yet, such processes have very rarely been modelled, as Okasha (2001, 2006) notes. Why is this so? We argue that just asserting group-level reproduction and the existence of group-level traits begs the question of how those properties arose. That is, how did the groups become cohesive enough for type 2 group selection to be operative? In some sense, in order for type 2 selection to occur the groups must posses a degree of individuality. We can see this by the fact that type 2 group selection is individual selection shifted one level up the biological hierarchy, so the groups are treated just like individuals. In general, for a type 2 model to be satisfying from an explanatory point of view, we need to be able to explain why we are interested in tracking the frequency of particular kinds of groups and not just individuals. That is, why are the groups cohesive enough to warrant this? This problem does not arise in type 1 models because there we are only interested in the frequency of individuals, and by definition what we call an individual is a cohesive unit, or a unit of evolution as Maynard Smith (1987) would call it. Interestingly, artificial group selection experiments (e.g., Wade 1976, 1977; Craig 1982; Muir 1996) are of the type 2 variety, but the problem of explaining cohesiveness does not arise there because the experimenter imposes individuality on the groups. That is, both group-level reproduction, and group fitness meaning number of descendant *groups*, are imposed as part of the experiment. These models show us that when these properties are imposed, there can be a response to selection on group-level traits. In the case of natural selection, however, we cannot rely on these properties being imposed. As we discuss in the next section, however, social niche construction can explain how these properties arise through the evolution of individual traits, and can thus bridge the gap between type 1 "individual selection in the context of groups", and type 2 "real group selection".

2.6.3 Social niche construction bridges the gap between individual- and group- selectionists

Proponents of an individual-level view to all social evolution, such as Grafen (1984), Dawkins (1994) and West et al. (2007a, 2008), see no need to invoke group selection in type 1 cases. This is because for them, it is just frequency-dependent individual selection in a structured population (Maynard Smith, 1998; Sterelny, 1996). On the other hand, although these authors would agree that type 2 models involve group selection, such models might seem unsatisfactory unless we can explain how the groups came to be cohesive targets of selection in their own right. As Dawkins would put it, why should we treat the groups, and not just the individuals, as vehicles?

Social niche construction, however, suggests that if the group structure is itself evolving in type 1 models, then it should not be regarded as merely a part of the environmental context of individual selection, in the way that abiotic environmental factors are. This is because individual adaptation can actually strengthen type 1 group selection, as we illustrate in this dissertation. Moreover, models that treat groups as merely a part of the selective environment cannot explain how those groups are themselves created. That is, type 1 group selection itself needs an evolutionary explanation, in the way that other static environmental contexts, such as aspects of the physical environment, do not. How does this connect with type 2 group selection? Social niche construction explains how individual adaptation reinforces group selection, by creating more cohesive groups. Such greater group cohesion could lead to type 2 group selection becoming operative on the groups. This suggests that the apparent impasse between type 1 and type 2 group selection can be bridged (see also Okasha 2006 and Michod 2005). Thus, social niche construction suggests that merely viewing the groups as part of the environmental context in type 1 scenarios is unsatisfactory, for doing so cannot explain how the groups were created and hence how group-level selection pressures on social behaviour came to be exerted. In addition, the evolution of type 1 group structure can potentially lead to "real" (*sensu* West et al. 2007a, 2008) type 2 group selection becoming operative, bridging the gap between these types of group selection.

Chapter 3

A Minimal Model of Social Niche Construction

In this chapter, we present a first investigation into how individuals can create selective environments favouring more or less cooperative behaviour, by modifying their population structure. Specifically, we consider how individuals can evolve a preference for the number of other individuals with which they found a group, for example, when colonising a resource patch. Such a preference directly affects the selection pressures on social behaviour, since groups founded by fewer individuals will tend to be less representative of the global population and more homogeneous. More formally, the groups' genotypic composition is a statistical sample of that of the global population, and taking a smaller sample size increases the between-sample variance and reduces within-sample variance. A smaller initial group size therefore tends to cluster individuals with the same social behaviour into groups together, providing a greater selection pressure for cooperation.

3.1 Introducing the aggregation and dispersal population structure

We consider a population structure based on the "Haystack model"[1] as conceived by Maynard Smith (1964), and later extended by Wilson (Wilson and Colwell, 1981; Wilson, 1987; Sober and Wilson, 1998). In this structure, groups of initial size n colonise a resource patch, and reproduce on that patch for a number of generations. After a fixed number of generations, the members of all groups disperse and freely-mix in a migrant pool, before new patches are again colonised by n individuals and the process repeats. Such a population structure can potentially allow cooperative behaviours to evolve, since groups founded by a greater proportion of cooperators will grow to a larger size, and

---

[1] The model is known as the Haystack model because Maynard Smith (1964) originally presented it as a model of an imaginary species of mouse that formed groups in discrete haystacks.

will hence contribute more individuals into the migrant pool at the dispersal stage. This is between-group selection of type 1. However, within-group selection also operates since cooperators decline in frequency *within* each group as reproduction and selection occurs. Whether or not cooperation increases in frequency *globally* then depends upon the balance of these two selective forces: the extra contribution of groups with more cooperators to the migrant pool must outweigh their decline within each group. An illustration of this population structure is shown in Figure 1.1.

Although somewhat idealised, this kind of population structure fits some organisms particularly well. In particular, organisms that live on an ephemeral resource patch where resource depletion triggers the dispersal stage. Examples include small arthropods (Wilson and Colwell, 1981), parasites (where the host is the ephemeral resource) (Wilson and Colwell, 1981; Bull et al., 1991; Williams and Nesse, 1991), and micro-colonies in bacterial biofilms (Ghannoum and O'Toole, 2004), or more generally any species that undergoes very large fluctuations in density due to population explosions and subsequent crashes (Bulmer and Taylor, 1980; Wilson and Colwell, 1981). However, our use of the model is motivated by the fact that it has few parameters and has been well studied in the social evolution literature (Bulmer and Taylor, 1980; Wilson and Colwell, 1981; Wilson, 1987; Bergstrom, 2002; Fletcher and Zwick, 2004). This means that it provides a simple and well understood population structure that we can extend by allowing, unlike in previous work, the structure to itself evolve. In the remainder of this dissertation we use the term “aggregation and dispersal”, rather than “Haystack”, model as this is a more descriptive term for the population structure. Specifically, “aggregation” refers to the individuals coming into groups, while “dispersal” refers to the individuals leaving those groups after a number of generations.

The parameters in the classical, fixed structure, aggregation and dispersal model that affect the evolution of cooperation are:

- the ratio of the group fitness benefit to the individual cost;
- the number of generations, t, spent reproducing within groups prior to the dispersal stage;
- the variance in initial group composition.

The greater the benefit-to-cost ratio of a cooperative act, the less stringent are the conditions for its evolution. Conversely, the greater the variance in the frequency of a cooperative allele between groups, the greater the range of cooperative behaviours that can evolve. In the limit, if there is no variance in cooperative behaviour between groups then the behaviour cannot evolve regardless of its benefit to others, since within-group selection favouring avoiding paying the immediate cost of cooperating will be the only selective force. At the other extreme, if all groups are clonal (with no new mutations occurring) then any cooperative behaviour can be selected for that has a benefit-to-cost ratio greater than 1, for in that case between-group selection will be the only selective force. This between-group variance will depend upon the initial variance in genotypic composition of the groups, which in turn depends on how groups are formed. As discussed above, if groups are formed randomly then a smaller initial group size will increase the between-group variance in the frequency of a cooperative allele, thus favouring the evolution of cooperation. Random group formation, however, represents something of a worst-case assumption (Wilson, 1977; Wilson and Dugatkin, 1997). Group formation could instead be assortative, such that individuals with the same social behaviour tend to join the same groups (Wilson and Dugatkin, 1997). This would again increase between-group variance and favour cooperation. Any such mechanism that increases between-group variance also conversely decreases within-group variance, and so reduces the scope for within-group selection favouring selfishness.

The number of generations spent within groups prior to dispersal can have two effects (Wilson, 1987). At the extreme, if the number of generations is large enough that the equilibrium allele frequencies are reached within each group, then the cooperative allele will be driven extinct in all groups that initially had at least one selfish allele present. This is because within-group selection is generally assumed to monotonically favour selfishness (an assumption that we show the implications of in the next chapters), and so causes the cooperative allele to decline in frequency within each mixed group with every generation. The assumption that each group must reach equilibrium allele frequencies before dispersal was made by Maynard Smith (1964) in his original presentation of the model (highlighted by Wilson 1987), where it was concluded that the likelihood of cooperation evolving in such a scenario was very bleak. However, the dispersal stage could happen before equilibrium allele frequencies are reached [2]. This could easily happen, for example, if each group lives on an ephemeral resource patch that becomes depleted after a few generations. Studies such as Wilson and Colwell (1981); Wilson (1987) have investigated the effects of this, and found that cooperation evolves far more readily if dispersal occurs before equilibrium. In some cases, an intermediate number of generations before dispersal can most favour the evolution of cooperation, since the initial between-group variance can become magnified over the first few generations of reproduction (Wilson, 1987; Fletcher and Zwick, 2004). On the other hand, in some cases only a single generation within groups can better favour cooperation, since this gives within-group selection fewer generations to act. This partly explains why the classic example of kin selection in sibling groups so effectively promotes cooperation (e.g., Williams and Williams 1957), since the siblings disperse immediately after the cooperative act towards each other has been expressed, reducing the effectiveness of opposing selection within the sibling group. Whether one or more generations within groups most favours cooperation depends upon the dynamics of within-group selection,

---

[2] In type 2 group selection, the analogue of dispersal is the formation and emigration of offspring propagules, that is, a subset of the group that breaks away after growth to occupy a new site.

which in turn depends on the type of the cooperative act, for example, whether it is weakly or strongly altruistic.

Although we have explained the above factors of population structure in terms of type 1 group selection theory, they can also easily be framed in terms of kin selection (for example, a version of the classical Haystack model where groups are founded by a single mated female and her clutch of offspring has been analysed in a kin selection framework by Taylor and Wilson 1988). To do so, we use the expanded version of Hamilton's rule introduced by Queller (1994) and advocated by West et al. (2002), which is able to deal with the effects of local competition between the beneficiaries of cooperation and hence with multiple generations of reproduction within groups before dispersal.

$$r_{xy}b - c - r_{xe}d > 0 \tag{3.1}$$

This version of Hamilton's rule is shown in equation 3.1 (West et al., 2002), where r_{xy} is the relatedness of the actor to the recipients of cooperation (the classical r term in the original version of Hamilton's rule), b is the benefit to the recipient, and c the cost to the actor. The extra term $r_{xe}d$ accounts for local competition. Specifically, d measures the extra competition due to increased crowding that is caused by cooperation raising the fitness of recipients and hence causing them to have more offspring. It is thus a measure of the negative density-dependent effects of extra growth. r_{xe} then measures the relatedness between actors and the individuals that suffer this increased crowding due to the cooperative act, and a reduction in this favours the evolution of cooperation. Equation 3.1 therefore highlights the need for two scales of interaction for cooperation to evolve: recipients of cooperation should disperse away from the actor to avoid ultimately competing with her for limited resources. This is therefore an explanation for why reproduction must not occur for too many generations before dispersal if cooperation is to evolve, and is equivalent to the notion of within-group selection opposing cooperation in the group selection framework. Similarly, the variance in initial group composition in the group selection framework is equivalent to r_{xy} at the time the groups are founded, once relatedness is measured as a statistical correlation between the allelic composition of group members relative to the mean allele frequencies in the global population (Grafen, 1984, 1985; Queller, 1994).

3.2 The evolution of the aggregation and dispersal structure

The basic aggregation and dispersal model, like nearly all models of social evolution, assumes that the parameters of the population structure remain fixed over evolutionary time. Thus, they do not consider what happens if these parameters can themselves be affected by individual traits and hence evolve in an adaptive manner. In particular, if the parameters can evolve, will they evolve towards conditions that favour cooperative or selfish behaviour? Here, we allow individual adaptation to determine whether the conditions for cooperation or selfishness are created.

To do so, we consider two competing initial group sizes, by giving individuals a genetic preference for living in groups founded by either many or few individuals. An individual's genotype thus contains two alleles. The first of these specifies the size of group that an individual will join from the migrant pool, the second the social behaviour of that individual. Selection for a small initial group size preference should favour cooperation, by in creasing between-group variance and the relatedness of the group members. On the other hand, selection for a large size preference should favour selfish behaviour, by reducing between-group variance and within-group relatedness.

We frame social behaviour in terms of resource utilisation strategy. Specifically, each group has a resource pool that is shared by all members. We assume that there is a trade-off between growth rate and yield, where the yield is the amount of biomass produced per unit of resource consumed. The cooperative strategy is then to reduce growth rate and increase yield, while the selfish strategy is to grow faster and hence reduce yield. If all individuals adopted the cooperative strategy, then the group would reach a larger total biomass. However, the selfish strategy will increase in frequency within each mixed group, since it has the higher growth rate, even though it will reduce the total biomass available from the resource. This therefore creates a "Tragedy of the Commons" scenario (Hardin, 1968), as discussed in Chapter 1, and the question is whether population structuring can avert this tragedy and allow a higher yield to evolve even when it is costly in terms of a reduced growth rate. This exact kind of trade-off between growth rate and yield has been demonstrated to occur in micro-organisms, for example, in ATP production as a result of metabolic trade-offs (Pfeiffer et al., 2001; MacLean, 2008).

3.3 A first model of the concurrent evolution of initial group size and social behaviour

The genotype of each individual in our model is haploid and specifies two binary traits:

1. Cooperative or selfish resource usage (specified as a growth rate, G_i, and a resource consumption rate C_i);
2. specification of the initial size of the group that that the individual will join (small or large).

We therefore consider 4 possible genotypes: cooperative + small, cooperative + large, selfish + small and selfish + large.

3.3.1 Reproduction within groups

Asexual reproduction occurs entirely within groups, as follows. Let n_i represent the (continuous) density of genotype i in a single group. In order to reproduce, these genotypes require a share of the group's resource influx, R. The size of this share that each genotype receives depends upon its growth and consumption rates relative to those of the other genotype(s), where the consumption rate should be understood as the amount of resource required for a genotype to make one copy of itself. The magnitude of the share of the total group resource that the genotype receives, r_i, is then as defined in Equation 3.2.

$$r_i = \frac{n_i G_i C_i}{\sum\limits_j (n_j G_j C_j)} R \tag{3.2}$$

Therefore, the genotype with the highest growth and consumption rates will receive the largest per capita share of the total resource. This means that the selfish genotype always receives more *per capita* resource than the cooperative type, and so will ultimately drive a cooperative genotype in the same group extinct at equilibrium.

In one version of the model, the amount of resource that a group receives at each time-step depends upon its size, with a larger *per capita* amount of resource allocated to larger groups. Specifically, groups of a large initial size receive 1.25 times more resource *per capita* than small groups. This implements a (weak) Allee effect (Allee, 1938; Odum and Allee, 1954; Avilés, 1999) whereby larger groups have an intrinsic advantage, for example, due to access to resources that a smaller group cannot obtain. Significantly, this provides an upwards pressure on initial group size away from that which would maximise cooperation and long term group productivity. We thus give in this version of the model the conditions that favour selfish behaviour an intrinsic advantage.

Given the resource share received by a genotype as described above, the density of the genotype in that group then changes according to Equation 3.3. The form of this equation is motivated as a simplification of existing models of bacterial colony growth (Pfeiffer et al., 2001; Pfeiffer and Bonhoeffer, 2003; Kreft, 2004). We allow the densities to be continuous numbers within groups.

$$n_i(t+1) = n_i(t) + \frac{r_i}{C_i} - D n_i(t) \tag{3.3}$$

As the difference equation (3.3) shows, the density of a genotype produced during reproduction depends on both growth (favouring the selfish type in the short-term) and consumption rates (favouring the cooperative type in the long-term under limited resource). The final term in (3.3) represents mortality, where D is a death rate that is constant to all genotypes.

3.3.2 The aggregation and dispersal process

The overall operation of our model is as stated in Algorithm 3.1.

Algorithm 3.1 Aggregation and dispersal model with two competing initial group sizes.

1. **Initialisation**: Initialise the migrant pool with a total density of N individuals, consisting of each of the 4 possible genotypes at equal frequency.
2. **Group formation (aggregation)**: Assign individuals in the migrant pool to groups, as described in the main text below.
3. **Reproduction**: Perform reproduction within groups for t generations, by performing t iterations of Equation 3.3 within each group.
4. **Migrant pool formation (dispersal)**: Return the progeny of each group to the migrant pool.
5. **Maintaining the global carrying capacity**: Rescale the migrant pool back to size N, retaining the proportion of individuals with each genotype, and rounding the density of each type to the nearest integer. This implements a global carrying capacity.
6. **Iteration**: Repeat from step 2 onwards for a number of cycles, T.

Assignment of genotype densities to groups occurs by the following process. Groups of the "small" initial size are created by choosing (without replacement) individuals at random who carry the small size allele on their genotype. Likewise, groups of the "large" initial size are created from randomly drawn individuals carrying the large allele. Therefore, although all individuals in a group will specify the same size trait, the composition of the group will be random with respect to the cooperative / selfish resource usage trait. This group formation process can be modelled by two hypergeometric distributions, one for individuals with the small size allele, and the other for individuals with the large allele. The parameters of these hypergeometric distributions are: N_h is the density of individuals with the size allele, n_h is the initial group size as specified by the size allele, and m_h is the frequency of the cooperative allele in the pool of individuals with that size allele (the subscript h denotes classical parameters of the hypergeometric distribution). Finally, if there are an insufficient number of individuals remaining in the migrant pool to form the appropriate group size, then those remaining unassigned to groups are discarded. Given a large population size, the effect of this is negligible.

3.4 Results: the concurrent evolution of initial group size and social behaviour

In this section, we first investigate the results of pairwise competition between genotypes, before proceeding to investigate the dynamics that occur when all 4 possible genotypes are present. Throughout this section, the parameter settings stated in Table 3.1 are used. We explore the effect of different initial group sizes and *per capita* resource allocations below.

Table 3.1: Parameter settings used throughout Chapter 3.

Parameter	**Value**
Growth rate (cooperative), G_c	0.018
Growth rate (selfish), G_s	0.02
Consumption rate (cooperative), C_c	0.1
Consumption rate (selfish), C_d	0.2
Small initial group size, n_{0s}	4
Large initial group size, n_{0l}	40
Total resource influx for small groups with Allee effect, R_s	4
Total resource influx for large groups with Allee effect, R_l	50
Migrant pool density, N	1000
Number of aggregation and dispersal cycles, T	4000

3.4.1 Preliminary Experimentation / Pairwise Competition

When deciding on values for “small” and “large” initial group sizes, it was essential that the imposition of the small size upon all members of the population lead to selection on behaviour favouring the cooperative trait, and the imposition of the large size lead to selection favouring the selfish trait. This would then give individuals the possibility of adaptively determining whether the environmental conditions, in the form of initial group size, lead to cooperation being selected for.

To determine suitable values for large and small initial group sizes, the 2-dimensional parameter space consisting of initial group size and the number of generations spent reproducing within groups prior to dispersal was sampled; all other parameters of the model were held constant. We set the resource influx, R, for each initial group size to be equal to that size, i.e., if the initial group size was 4, then the resource influx per generation was also 4. Thus, the initial *per capita* resource influx across all group sizes was unity. Figure 3.1 plots the 2-dimensional parameter space with respect to whether cooperation was selected for. We report cooperation as being selected for only if the cooperative trait reached fixation in the global population at equilibrium, i.e., after 4000 aggregation and dispersal cycles.

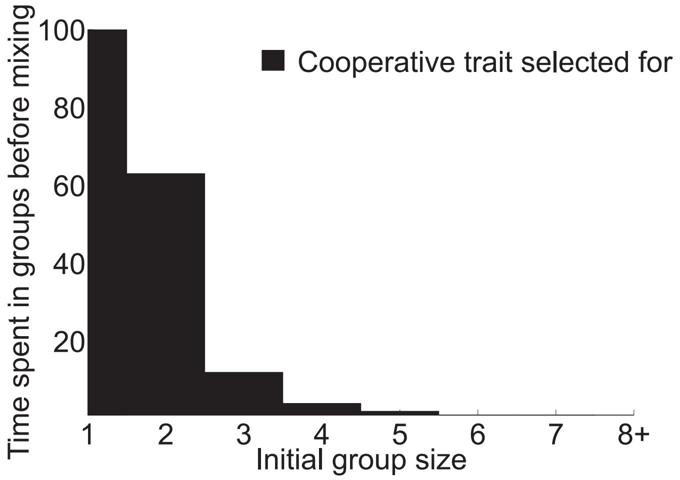


Figure 3.1: Exploration of the 2-dimensional parameter space over initial group size and the number of generations within groups before disperal.

As expected from the classical theory discussed at the start of this chapter, this sample of the space shows that cooperation is favoured by a population structure where groups are founded by few individuals and where a short number of generations is spent within those groups prior to dispersal and global mixing. A small initial group size favours cooperation by increasing the variance in group composition, thereby strengthening selection acting between groups, i.e. selection that is due to the difference in group sizes after reproduction Sober and Wilson (1998). Equivalently, they increase the relatedness between a cooperator and the other group members that benefit from her reduced growth rate. A small number of generations within groups prior to mixing favours cooperation since the longer the time spent in groups, the greater the decline in frequency of cooperators within all groups that also contain cheaters (Sober and Wilson, 1998). From the kin selection perspective, it decreases the relatedness between the actor and the other individuals who suffer the effects of crowding due to the increased fecundity of the recipient of cooperation, since the groups are reformed more frequently.

Given the results in Figure 3.1, we set the small group size at 4 and the large at 40, with the number of generations within groups fixed at 4. This then creates the situation in which cooperative + small outcompetes selfish + small, and selfish + large outcompetes cooperative + large. Therefore, our definition of small initial size favours cooperative resource usage, while our definition of large favours selfish usage.

3.4.2 Allowing Both Social Behaviour and Population Structure to Evolve

We initially considered the case where there is no intrinsic advantage to larger groups, thus groups of the small initial size received 4 units of resource per generation, and groups of the large initial size 40. We competed cooperative + small against selfish + large by starting each genotype in equal frequency in the global population. We found that cooperative + small always outcompeted selfish + large; the results of an illustrative run are shown in Figure 3.2 (all runs showed the same qualitative trends, and the standard error was close to 0. This is due to the fact that everything in the model is a deterministic process, apart from group formation). Both the cooperative + small and selfish + small genotypes initially increased in frequency, meaning that the small groups population strucutre was initially favoured by both types. In particular, selfish individuals did well in the small groups for a short time, since they were able to exploit the cooperators in those groups. However, this could not be sustained, because we know from Figure 3.1 that the selfish type is ultimately driven extinct by group selection in such small groups. In conclusion, even though a population structure favouring selfishness was available, this was never selected for, since both cooperators and selfish individuals (initially) do better in the smaller groups where there is a greater frequency of cooperation. As this group size becomes dominant, cooperative behaviour sweeps to fixation.

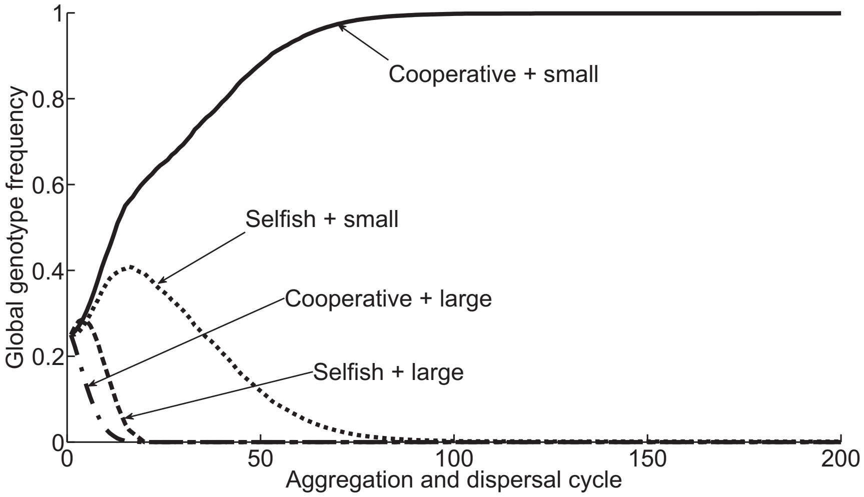


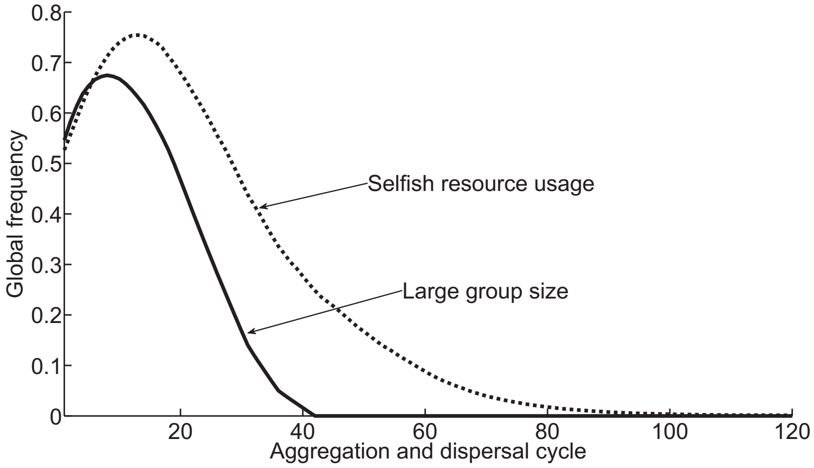
Figure 3.2: Evolution of population structure and social behaviour with no intrinsic advantage to larger groups.

We can provide a logical argument for why, in the absence of any other selection pressures on population structure, this must be the case. Cooperation increases in frequency in groups of a smaller initial size, compared to those of a larger initial size. This follows immediately from the fact that the between-group variance and relatedness is higher in the smaller groups, and so cooperators in such groups tend to interact more with each other and reduce exploitative interactions with selfish individuals (the “small” size was chosen so that it would provide sufficient extra between-group variance to select for greater cooperation, as shown in Section 3.4.1). Thus, greater cooperation is selected for in the smaller groups. If the individuals in such groups are those with a genetic preference for being in them, i.e., they posses the small size allele, then linkage disequilibrium will be generated between the cooperation and small size alleles. As a result bearers of the small allele will receive a greater *per capita* share of the benefits of cooperation, and so will be fitter than bearers of the large allele, all other factors being equal. Selection for cooperative behaviour that occurs in small groups therefore drives selection for those small groups themselves, initiating a positive feedback process (Breden and Wade, 1991; Crespi, 2004) that drives the cooperative + small genotype to fixation. We can conclude that any component of selection on population structure that is due to social behaviour must be in the direction of a structure that supports cooperation. This is because, with all other factors held constant, selection towards a structure favouring increased selfishness would reduce the fitness of individuals in that structure. This follows from the definition of cooperative and selfish behaviour, and the fact that linkage disequilibrium is generated between the behavioural and structural alleles by the fact that individuals with a preference for a certain structure find themselves living in that structure more frequently than individuals that do not have the preference (in this model, the match between preference and the structure inhabited is perfect). Thus, if individuals have an allele that determines which structure they live in, those choosing the selfish structure would be less fit, and so an allele that creates this structure must decline in frequency.

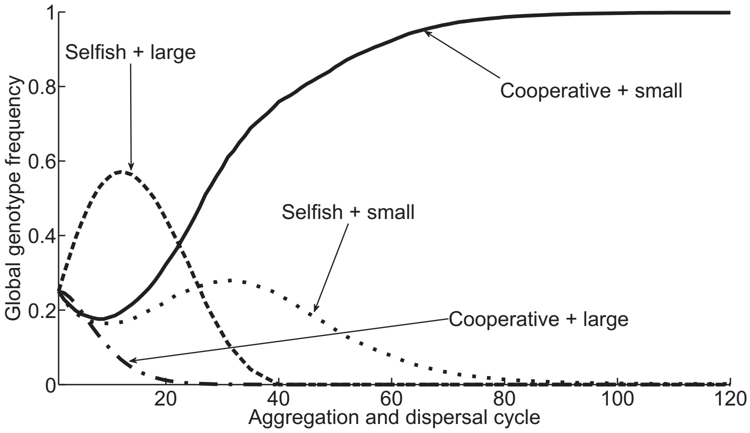
The question that we turn to next is, what happens if there are other components of selection on population structure apart from those resulting from social behaviour? In particular, we consider an Allee effect that provides an intrinsic advantage to groups founded by more individuals. This thus provides a component of selection on initial group size that is in the opposite direction to that caused by social behaviour. How will these components interplay to form the resultant vector of selection on initial group size? As discussed previously, we incorporate an Allee effect by giving small groups 4 units of resource influx per generation, and large groups 50, i.e, 1.25 times more *per capita*.

Because of the Allee effect, preliminary results showed that in pairwise competition, cooperative + small is outcompeted by cooperative + large. This is due to the greater *per capita* resource influx provided to large groups, which allows such groups to grow to a larger size, assuming all other conditions are the same. Likewise, selfish + large outcompetes selfish + small. Given that large size preference genotypes outcompete their small counterparts, and that selfish resource usage is favoured in a population structure where groups are of a large initial size, it might be expected that the population would evolve towards a state with the selfish + large genotype at fixation. In particular, selection on the size allele due to the Allee effect is direct, for the benefits of a larger size are immediate. On the other hand, the component of selection on initial size due to social behaviour is indirect, for it occurs only after cooperation has increased in frequency within those groups. We might therefore expect the direct component of selection on population structure to outweigh the indirect, and so the conditions favouring selfishness would be created after all.

In order to investigate this, we started each of the 4 possible genotypes in equal frequency in the migrant pool, and recorded the change in genotype frequencies over 4000 aggregation and dispersal cycles, which was a sufficient length of time for a global equilibrium to be reached. We show results from a single run here in order to highlight the transient dynamics; there was no qualitative variation in these dynamics over 30 runs. This is because the only source of stochasticity in our model is in the group formation stage, but over many aggregation and dispersal cycles this does not have a significant effect, and can be modelled by the hypergoemetric distribution as described in Section 3.3.2.



(**A**)



(**B**)

Figure 3.3: A) average population structure and social behaviour through time. B) change in genotype frequencies over time.

Figure **3.3A** shows how the average population structure, in terms of initial group size, and the average social behaviour, in terms of resource usage strategy, change over time. The key point is that at the start neither small groups nor cooperators are favoured; this is shown by an increase in the frequency of both large groups and selfish cheats. Therefore, since the initial conditions favour both large groups and selfish individuals, it might be expected that the equilibrium reached would consist entirely of large groups of the selfish type. However, this is not what occurs. Instead, after around 20 aggregation and dispersal cycles, both small groups and cooperators become favoured by selection. Therefore, the selective pressures change in a way that was not easy to predict from the pairwise competition of genotypes alone.

Figure **3.3B** helps to explain why this occurs, by showing how all possible combinations of population structure and social behaviour change over time. The figure shows that initially, the population structure of large groups is favoured; this is due to their per capita resource advantage. In such large groups, the selfish inefficient resource usage strategy is favoured, as previously discussed.

However, the selfish + large genotype does not reach fixation. This is because selfish cheats benefit from consuming resource in mixed groups with cooperators, since they can consume the left-over resource that follows from the cooperators efficiency. As the selfish + large type increases in frequency, it reduces the frequency of cooperative + large and hence the number of cooperators that it can exploit. However, there are still cooperative + small individuals present in the population. This then creates a selective advantage for the cheats that are in small groups rather than large, because the small groups still contain cooperators that can be exploited. Thus, cheats change their preference to create small groups rather than large. This explains the increase in frequency of both small types after around 10 aggregation and dispersal cycles. By changing size preference, however, the cheats create the very population structure in which cooperation thrives. Thus, once small becomes the dominant size allele, cooperative social behaviour fixes in the population. In particular, we know that cooperators drive selfish cheats extinct where the initial group size is 4 and $t = 4$ (as shown in Figure 3.1), and this is the resulting state when cheats change their initial size preference from large to small.

3.5 Discussion

We have developed here a minimal model of social niche construction, by considering competition between two initial group sizes. The “small” initial group size allele creates a population structure that selects for cooperative behaviour, by increasing the between-group variance / relatedness of the group members. Conversely, the “large” allele creates a population structure that selects for selfish behaviour, by increasing *within*-group variance and hence the strength of within-group selection. Individuals with the “small” size allele will therefore evolve to invest in between-group competition, whereas those with the "large" allele will evolve to invest in individual competition within their group. On the face of it, it may not seem obvious which of these strategies would be favoured. Indeed, one may think that individuals would evolve to invest in individual competition within their group, since the benefits of this would be immediate, whereas the benefits of investing in between-group competition would only be realised when the groups mix and are reformed. Under natural selection, we should expect social niche constructors to create an environment that favours themselves compared to other individuals, for this is the only way that the social niche constructing trait could be selected for. If we equate individual competition with within-group competition, then we would expect social niche constructors would create an environment that favours themselves in within-group competition with other individuals.

However, our model has provided an illustration that this is in fact not what happens when a niche constructing trait affects population structure and hence the evolution of social behaviour. Consider first the case where there is no Allee effect, and so the only pressures on group size come from social behaviour. To understand why small, cooperative groups are favoured, we need to take account of the fact that group size preference effectively partitions the population into two. Specifically, a group size preference allele must mean that individuals with a certain size preference tend to find themselves living in a group of that size more frequently than individuals that do not have that preference. If this were not the case, then it would not really be an allele that affected the group size of its bearers. In this model, the match between preference and size of group inhabited is perfect. This then partitions the population into two: individuals with the small preference allele live and interact in small groups, individuals with the large allele live in large groups. If small groups are more affected by group selection than large, then the mean fitness of individuals in small groups will be greater. This is because (type 1) group selection by definition favours traits that increase the mean fitness of group members, for this is what we mean by group fitness in a type 1 scenario. Thus, over a number of aggregation and dispersal cycles, cooperative traits that increase the mean fitness of all individuals will be selected for in the "small" subpopulation. Competition between the two subpopulations occurs when the migrant pool is formed, since individuals from all groups, regardless of size, compete to form the next generation of groups. In particular, when the migrant pool is formed, if individuals in small groups are on average fitter, then they will make up a larger frequency of the migrant pool and hence increase in frequency in the global population. Since individuals from small groups carry the small preference allele, both the small and cooperative alleles increase in frequency together. Put another way, linkage disequilibrium evolves between group size and behavioural alleles, and the end result shows this in extreme form by the fixation of the cooperative + small genotype.

In summary, social niche construction leads to evolution of small initial group size, and cooperation, in this model because:

1. Small initial group size in isolation selects for cooperation, large initial group size does not (Figure 3.1).
2. Cooperative behaviour raises the mean fitness of group members (by definition).
3. The size allele partitions the population into those living in small or large groups (definition of size preference + assumption that match between preference and actual group size is exact).
4. Because of points 1 and 3, linkage disequilibrium evolves between the size preference and socio-behavioural alleles. In particular, only two genotypes will be supported: cooperative + small and selfish + large.
5. Because of point 2, individuals with the cooperative + small genotype outcompete those with the selfish + large in the migrant pool formation stage, and hence increase in frequency in the global population.

This process relies on a matching between group size preference and the group size inhabited, and on greater cooperation being selected for in small groups than large. This second factor turns out to be critical, as we will demonstrate in Chapter 6.

The above argument assumed that the only pressures on initial group size came from social behaviour, as in the first version of our model. When there are other factors, such as an Allee effect favouring larger groups, then matters are more complicated, as the non-monotonic selective dynamics in Figure 3.3**B** illustrate. However, we can still see some of the same principles at work. In particular, the evolution of linkage disequilibrium between size and behavioural alleles again takes centre stage in the story. This time, the intrinsic advantage to initially larger groups, in the form of a greater *per capita* resource allocation, means that large groups are initially selectively favoured, despite the increased cooperation in small groups. However, this advantage is frequency dependent, and only occurs when the cooperative + large genotype is present in appreciable frequency. Essentially, when both cooperative + large and selfish + large genotypes are present, the large allele enjoys the benefits both of cooperation and of greater resource access. It is thus fitter than the small allele, which only experiences the benefit of cooperation. This advantage is not sustainable, however, because a population structure of large groups does not support cooperation (see Figure 3.1), and so the cooperative + large genotype must decrease in frequency, eventually to extinction. We can see from the grey dotted line in Figure 3.3**B** that this trend of decreasing frequency of cooperative + large is monotonic from the start. As the cooperative + large genotype reaches extinction, the large allele no longer enjoys the benefit of cooperation. It was then the case with the model parameters examined that cooperation yielded a greater *per capita* fitness than did the intrinsic benefits of large size, and so the cooperative + small genotype fixed in the population. Indeed, this is the only interesting case, for otherwise cooperation could never be selectively advantageous. What is then interesting is the conditions under which the fitter state of cooperation can be reached by a population that modifies its own structure, for the fact that one state is fitter than another does not entail that that state will be reached. This model illustrates that when structure is evolving, the cooperative state can be reached even when cooperation is initially being selected *against*. Significantly, the model illustrates that it is possible for the component of selection on population structure due to social behaviour to overcome another opposing component, even when that opposing component is due to an immediate fitness benefit in the form of a greater *per capita* resource allocation. The model also illustrates that the interplay between these opposing components of selection on the group size allele can be non-monotonic, and that the transient behaviour in the first few generations is not a good predictor of the end state.

We have developed this model as a minimal illustration of social niche construction. To keep the model minimal, we have only considered competition between two group size alleles, both of which were present in the initial population. When combined with the socio-behavioural alleles, this produced 4 possible genotypes whose frequency we tracked. The presence of one group size that favours cooperation, and another that favours selfishness, provides the minimal conditions for the evolution of population structure to be affected by social behaviour. We were thus able to explore whether individual selection on group size would create the conditions for cooperation or selfishness.

However, the scenario that we would ultimately like to consider is one where a population starts out living in a population structure where selfish behaviour dominates, and then evolves by a series of mutations to create the conditions for cooperation. For example, we would like to start the population out fixed for one large group size allele, and then investigate the dynamics that occur when both slightly smaller and slightly larger size alleles arise by mutation. For selection to lead to a group structure favouring cooperation, it must be the case that moving to a slightly smaller group size increases the degree of cooperation experienced by individuals living in those slightly smaller groups. If this does not occur, then individuals living in the smaller groups will not be fitter than those in the larger ones, all other factors being equal. As a result, the small size allele would not be selectively advantageous and so would not be able to invade a population fixed for the large allele. That is, there must be an adaptive *gradient* on the size allele such that a small decrease, as would be plausible by mutation, would increase the fitness of its carriers. In the absence of any other ecological advantage to smaller groups, this fitness increase must come from a change in the selection pressures on social behaviour.

Yet, Figure 3.1 demonstrates that such a fitness increase would not occur over a large range of initial conditions in this model. In particular, above an initial group size of 6, decreasing initial group size by 1 does not yield any increase in cooperation. Thus, if the population were fixed for a size allele of 7 or greater, a mutant allele of 1 size smaller could not invade. Thus, over much of the parameter space the conditions for cooperation could not be created by mutation and selection, unless the mutations were very large. The problem is that as the initial group size increases, any effect of group selection on social behaviour rapidly reaches zero, as shown in Figure 3.1. Because of this, moving to a slightly smaller initial group size does not increase the effect of group selection, and hence does not increase cooperation and the fitness of individuals living in those groups.

Must it always be the case that any effect at all of group selection on cooperation rapidly reaches zero as the initial group size increases? In the next two chapters, we will show that this effect rests on an assumption of competitive exclusion within groups. In particular the model here, like nearly all other models of group selection, assumes that within-group selection on social behaviour is directional and always favours the selfish allele. Thus, in the limit, cooperation would be driven extinct within each group. However, there are many cases where a stable polymorphism of social types could be maintained in each group by negative frequency-dependent selection. In Chapter 4, we examine this dynamic in a two-species group, where one (selfish) species dominates within the group and lowers the effective carrying capacity, but does not drive the other more cooperative species extinct, because their niches only partially overlap. In Chapter 5, we examine negative frequency-dependent selection on cooperation in the context of the Snowdrift game (Sugden, 1986; Doebeli and Hauert, 2005). In general, we find that where negative frequency-dependent selection operates within groups (of single or multiple species), then the effect of group selection does not rapidly tail to zero. Rather, a small decrease in initial group size can increase the effect of group selection, and hence the frequency of cooperation, over a much greater range of parameters. Thus, an adaptive gradient on a population structure-modifying allele is present from a much larger set of conditions. This means that where selection on social behaviour exhibits negative frequency-dependence, the adaptive evolution of population structure in support of greater cooperation can occur from a much larger range of initial conditions, as we demonstrate in Chapter 6.

Chapter 4

Behavioural Polymorphism and Group Selection

4.1 Introduction

Natural selection at any level of the biological hierarchy requires there to be a variance in fitness at that level (Price, 1972; Sober and Wilson, 1998; Michod, 1999; Okasha, 2006). In models of group selection, this variance is usually generated through the stochastic processes of either genetic drift (e.g., Wright 1945; Levin and Kilmer 1974) or group formation through random sampling from the global population (e.g., Wilson 1975a, 1987; Fletcher and Zwick 2004). In both of these cases, the group size must be small in order for the stochastic process to reliably create a significant level of between-group variance. Consequently, it is commonly held that group selection can only be effective in groups of an implausibly small size.

Proponents of group selection sometimes point out that significant levels of between-group variance can be generated in even large sized groups if group formation is an assortative, rather than purely random, process. That is, if individuals of the same type tend to somehow find themselves in groups with each other then the between-group variance will be larger. If this is the case then compared to the random group formation case more of the variance in fitness will be at the group, rather than the individual, level, and so one might expect to see an effect of group selection where there is none under random group composition. This could then provide a way out of the problem discussed at the end of the previous chapter, i.e., that the lack of *any* effect of group selection over a large range of group sizes entails that there is no selective gradient over which the niche constructing dynamic can operate. Assortative grouping, then, could be expected to allow an effect of group selection to be seen over a much larger range of group sizes, thereby providing a selective gradient towards smaller groups and increased cooperation to follow.

However, although such assortative grouping can in principle result from a variety of mechanisms (e.g., Wilson and Dugatkin 1997), it is commonly believed that kin grouping, where the group members are related by descent from a common ancestor and are hence more similar to each other than to members of other groups, is by far the most common way in which positive assortment is achieved in nature. Consequently, such processes are commonly said to operate over a large range of parameter settings precisely because they are kin, rather than group, selection[1] (e.g., Maynard Smith 1964, 1976; Grafen 1984). We consider here, however, a different reason as to why group selection may operate over a larger range of group sizes than is commonly believed. Specifically, we consider the within-group selective dynamics of groups of initially *random* composition, but investigate the effect of different types of within-group selection regimes.

Traditionally, it is assumed that within-group selection will lead to the competitive exclusion of a cooperative type by its more selfish counterpart, that is, that within-group selection is purely directional. Is such competitive exclusion always inevitable within social groups? We would argue that the answer is no. For example, cooperative and selfish behaviours can be subject to negative frequency-dependent selection, leading to a stable polymorphism of cooperative and selfish types within groups (as can be modelled by the Snowdrift or Hawk-Dove game (Maynard Smith, 1982; Sugden, 1986; Doebeli and Hauert, 2005), see Chapter 5). Such polymorphisms can occur even in the absence of group selection, if the benefits of cooperation change with the frequency of cooperators in the group (Dugatkin et al., 2003; Doebeli and Hauert, 2005; Gore et al., 2009). Examples of where cooperators and cheats coexist within single trait groups, due to such negative frequency-dependent selection, include cooperative foraging in spiders (Pruitt and Riechert, 2009) and extracellular enzyme production in yeast (Gore et al., 2009) (Chapter 5 provides further examples). Although these are examples where the groups consist of a single species, a coexistence of different species could also potentially occur within a trait group. For example, bacterial biofilms are commonly composed of multiple coexisting species (Watnick and Kolter, 2000), and cooperative interactions between these species have been implicated in the high resistance of biofilms to antimicrobial agents (Burmolle et al., 2006).

Results presented in this chapter show that measurable group selection effects are sustained over a much larger range of group sizes, in cases where a within-group stable coexistence of social types exists. This can potentially overcome the difficulty discussed at the end of Chapter 3, i.e., that no effect of group selection is seen at all over a wide range of parameter space under competitive exclusion dynamics. Thus, coexistence dynamics might provide an individual adaptive gradient towards smaller groups that enhance group selection over a much larger range of parameter space, and hence

---

[1] Kin selection here is used in the older sense of relatedness due to shared ancestry, rather than in the modern sense of any positive type correlation, since if the relatedness is not by descent then it seems hard to argue against describing the model as one of type 1 group selection. It is this older sense of the word that Maynard Smith and Grafen seemed to use.

allow social niche construction to occur from a much larger range of initial conditions. The fact that our model considers groups of initially random composition (rather than formed assortatively) is very important when considering the plausibility of the process, since random group formation represents the most conservative assumption of the level of variance that may initially exist between groups in natural populations (Wilson, 1980).

The model presented herein uses the aggregation and dispersal process from the previous chapter, but considers different types of within-group selection, as discussed in the next section.

4.2 Modelling within-group dynamics: competitive exclusion versus coexistence

Classical models of group selection consider a scenario where a selfish type ultimately drives its cooperative counterpart extinct at within-group equilibrium. In particular, fitness functions of the following form, first proposed by Wright (1945) but subsequently used in a plethora of other models (Williams and Williams, 1957; Maynard Smith, 1964; Charnov and Krebs, 1975; Wilson, 1980, 1987), are typically used to model within-group selection:

$$w_s = 1 + pb \qquad\qquad w_c = (1 + pb)(1 - a) \tag{4.1}$$

In the above equations, w_s and w_c denote the per capita fitness of selfish and cooperative individuals within a group, respectively. Cooperators, whose proportion within the group is denoted by p, confer a fitness benefit b on every group member. Crucially, both types receive this benefit, while only cooperators pay a cost, represented by the selection coefficient against cooperation, a. It is then clear that if these equations are iterated until equilibrium is reached then the selfish type will be driven to fixation within a single group (in the absence of any group selection effects). Thus, such equations tacitly assume competitive exclusion of social behaviours.

Both competitive exclusion and stable coexistence within-group dynamics can instead be modelled using the standard two-species symmetric Lotka-Volterra competition equations (e.g. May, 1976), which we use here as a well studied didactic example of coexistence and competitive exclusion. For implementation purposes, we use the following difference equation as a discrete approximation:

$$N_{i(t+1)} = N_{i(t)} + \left[M_i N_i \left(\frac{K_i - \alpha_{ii} N_i - \alpha_{ij} N_j}{K_i} \right) \right]. \tag{4.2}$$

In the above equation, $N_{i(t)}$ is the biomass of species (type) i at time t, and M_i the intrinsic per capita growth rate. Each species has an intrinsic carrying capacity K_i, which is then modified through interspecific interactions. Specifically, the per capita effect of species j on species i is given by α_{ij}, the coefficient of interaction. All such interactions are competitive in the above equation, as ensured by the negative sign and the stipulation that all $\alpha > 0$. Similarly, α_{ii} denotes the negative density-dependent effect of species i on itself that prevents unbounded exponential growth. This coefficient can be seen as representing crowding and can vary for different species.

We define selfish (s) and cooperative (c) strategies in the above equation through settings of the within- and between-type interaction coefficients. Specifically, a selfish type is defined as having a large negative per capita effect on both itself (α_{ss}) and the other type (α_{cs}). A cooperative type is then defined as having correspondingly smaller per capita negative effects (α_{cc} and α_{sc}). A pure group of cooperators will therefore grow to a larger size than a pure group of selfish individuals, creating a group productivity differential on which selection can potentially act. However, within mixed-groups the selfish type will reach the larger frequency (provided that α_{ss} is not too large), since $\alpha_{cs} > \alpha_{sc}$. In other words, cooperators are favoured by between-group selection, while selfish individuals are favoured under within-group selection, exactly as in a classic group selection scenario. It should be noted that our definition of cooperative behaviour corresponds to weak, rather than strong, altruism (Wilson, 1980). This follows because although cooperation confers a relative fitness disadvantage compared to a selfish individual within the same group, it nevertheless increases the absolute fitness of all group members, including the cooperator.

It is well known that a stable coexistence of both types occurs in such a model when competition for resources (space, food, etc.) between individuals of the *same* type is stronger than competition between individuals of different types. Such a case corresponds at the ecological level to species occupying different niches, i.e. only partially overlapping in their resource requirements (May, 1976). Conversely, if between-type competition is stronger than within-type competition then competitive exclusion of one type will occur. Between- and within-type competition are both modelled in the Lotka-Volterra equations through the settings of the interaction coefficients. Throughout this chapter, we assume the following:

1. $\alpha_{cc} < \alpha_{ss}$ and $\alpha_{sc} < \alpha_{cs}$, i.e. that cooperators have lower negative density-dependent effects on themselves and others;
2. $\alpha_{cs}\alpha_{sc} \leq 1$;
3. M and K, the intrinsic per capita growth rates and carrying capacities respectively, are the same for both types.

Given these assumptions, competitive exclusion of the cooperative type occurs when $\alpha_{ss} < \alpha_{cs}$, producing qualitatively similar dynamics to those of the traditional within-group selection equations described in Section 4.2. However, when $\alpha_{ss} > \alpha_{cs}$ then the cooperative type is maintained at within-group equilibrium at an above-zero frequency, i.e. a stable coexistence of types occurs. When the interaction coefficients of the cooperative strategy are fixed, the equilibrium frequency at which it is maintained then depends upon the settings of α_{ss} and α_{cs}, i.e. the magnitude of the negative density-dependent effects of selfish individuals on themselves and cooperators, respectively. In addition, the within-group equilibrium is reached more quickly the greater the effects of the selfish type. Although a Lotka-Volterra model such as this is typically interpreted at the ecological level as representing species interactions, it could also be interpreted as a model of allelic competition dynamics within a single species group. In particular, coexistence Lotka-Volterra dynamics are analogous to negative frequency-dependent selection leading to a stable allelic polymorphism within a group. Conversely, competitive exclusion of one species by another is analogous to directional selection driving one allele to fixation. The motivation for using the language of allelic competition is to facilitate comparison with classical group selection models, which consider competition between selfish and cooperative alleles.

Our use of the Lotka-Volterra equations in the model should be contrasted from their use in community or ecosystem selection models (Wilson, 1992; Penn, 2003, 2006). Such models do not consider explicit cooperative and selfish types in the fashion of traditional group selection models. Instead, they examine the complex within-group dynamics that arise when a larger number of types are present. These complex dynamics can give rise to multiple within-group attractors, which can then provide a source of variation in their own right upon which selection can act (Penn, 2003). By contrast, this model considers simple two-type within-group dynamics, where only a single group attractor exists (either coexistence or competitive exclusion, as discussed above). As far as we are aware, our use of the Lotka-Volterra equations to define explicit selfish and cooperative strategies is novel.

4.3 Results

The parameter settings used for the Lotka-Volterra equations throughout this chapter are shown in Table 4.1. Changing between competitive exclusion and coexistence within-group dynamics is achieved by simply switching over the values of α_{ss} and α_{cs}, since that determines whether $\alpha_{ss} < \alpha_{cs}$ and hence whether competitive exclusion occurs. The values of the interaction coefficients in Table 4.1 produce representative within- and between-group dynamics; other settings produce the same qualitative trends. In this section, we first present results using classical competitive exclusion dynamics, and then contrast these to results from the coexistence case.

Parameter	**Value (competitive exclusion)**	**Value (coexistance)**
α_{ss}	**1.9**	**2**
α_{cs}	**2**	**1.9**
α_{cc}	1	1
α_{sc}	0.5	0.5
K	100	100
M	0.1	0.1

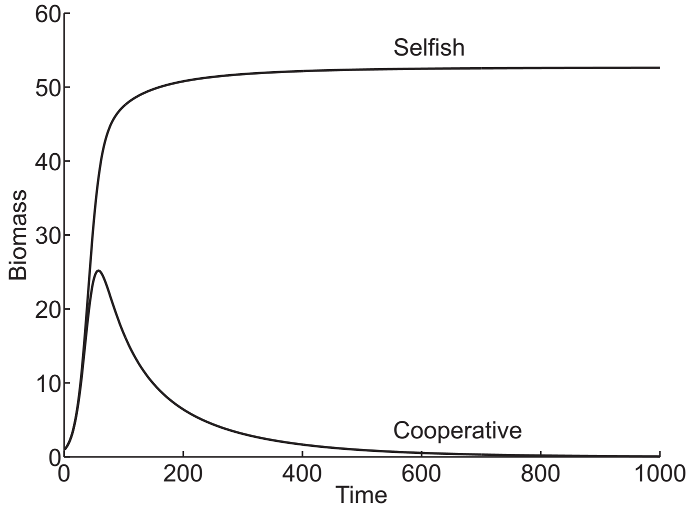
Table 4.1: Parameter settings of the Lotka-Volterra equation. Note that the only difference between the competitive exclusion and coexistence settings is a swapping of the values of α_{ss} and α_{cs}.

4.3.1 Group Selection Dynamics in the Competitive Exclusion Case

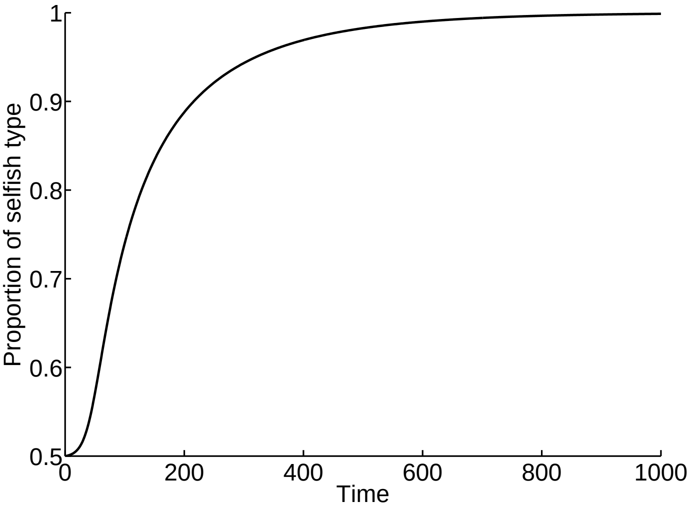
The within-group dynamics for a group initialised with unit biomass of each type are shown in Figures 4.1**A** and 4.1**B**, for the competitive exclusion case. Initially, both types are in their growth phase; their biomass is below the intrinsic type carrying capacity of 100. However, the selfish type grows at a faster rate, despite the fact that their intrinsic growth rates, M, are the same. This is because of the greater negative density-dependent effect of the selfish type on cooperators, i.e. $\alpha_{cs} > \alpha_{sc}$. Finally, since $\alpha_{cs} > \alpha_{ss}$, the cooperative type is driven to extinction. Furthermore, as Figure 4.1**B** shows, the proportion of selfish individuals increases monotonically. Such behaviour is qualitatively identical to that of directional within-group selection for a selfish allele in classical group selection models (e.g. Wright 1945; Wilson 1980).

Now let us consider global dynamics under group selection in this competitive exclusion case. In order for group selection to operate through an aggregation and dispersal process, a difference in group size at the dispersal stage must exist. Figure 4.2 illustrates how final group size varies as a function of the time spent in the group prior to dispersal, for various starting frequencies of cooperators in groups of initial size 10. It can be seen from this graph that, using the parameters described in Table 4.1, groups with a greater proportion of cooperators do indeed grow to a larger size. In addition, the results for the coexistence case, where α_{ss} and α_{cs} are swapped, are quantitatively similar. These results therefore confirm that group selection can in principle operate, since there is a variation in group productivity on which selection can act.

To determine the magnitude of the effect of group selection, the aggregation and dispersal process was executed for 5000 iterations, which preliminary experimentation had shown to be a sufficient length of time for a global equilibrium to be reached, using the within-group parameter settings described in Table 4.1. Equation 4.2 was iterated 30 times in the reproduction stage, while the global population size was maintained at 5000. Initial group size was then varied from 1 to 100 inclusive, while the migrant pool was initialised with 50% of each type. The result of this process after 5000 aggregation and dispersal cycles is shown in Figure 4.3, where 'effect of group selection' on the



(**A**)



(**B**)

Figure 4.1: Competitive exclusion within-group dynamics. A)Biomass of each type. B) Proportion of selfish type.

y-axis is defined as the difference between the frequency of the selfish type at within-group equilibrium and the global frequency of the selfish type after 5000 aggregation and dispersal cycles. Since the within-group equilibrium in the competitive exclusion case is the selfish type at 100%, the y-axis equivilantely shows the global frequency of cooperators in this case. Furthermore, it should also be stressed that the within-group equilibrium is the equilibrium that would be reached in an unstructured population where there were no groups. The y-axis therefore shows the effect that group structure is having on the outcome of evolution compared to that in an unstructured population.

There are two points to note from Figure 4.3. Firstly, increasing the initial group size decreases the effect of group selection, and consequently the global proportion of

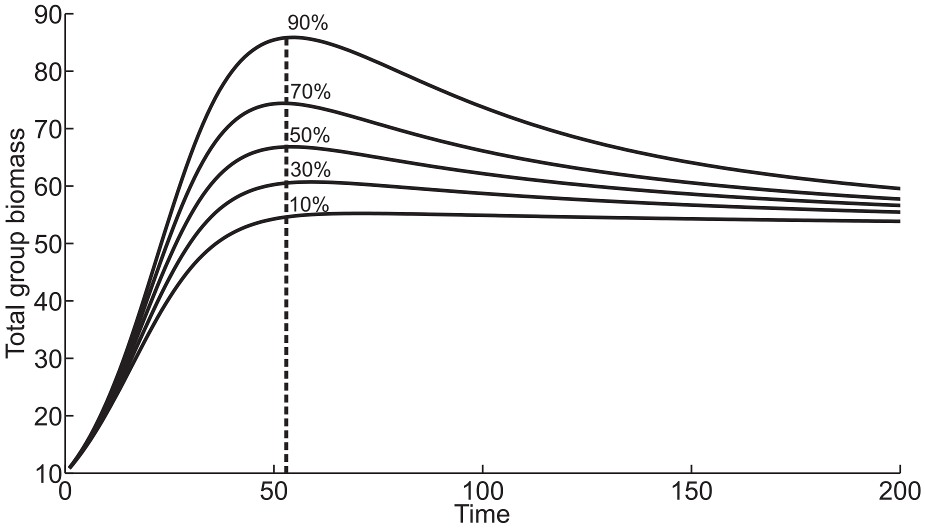


Figure 4.2: Final group size as a function of time spent reproducing within groups; initial group size 10 with various % of cooperators (competitive exclusion case shown; coexistence quantitatively similar). Dotted line shows time at which difference in group productivity is greatest.

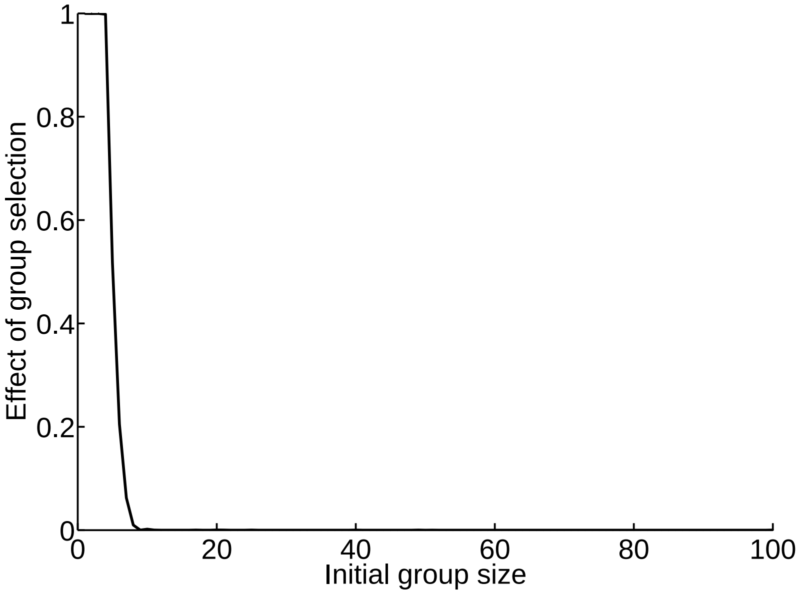


Figure 4.3: 'Effect of group selection' (see text) as a function of initial group size in the competitive exclusion case.

cooperators. In particular, for small group sizes, the cooperative type reaches global fixation (and remains there because we do not reintroduce types by mutation). However, for group sizes above 10, it is driven extinct. This follows from the fact that the between-group variance necessary for group selection to act is generated by random sampling from the migrant pool, and therefore rests on the existence of a small initial group size, as previously discussed.

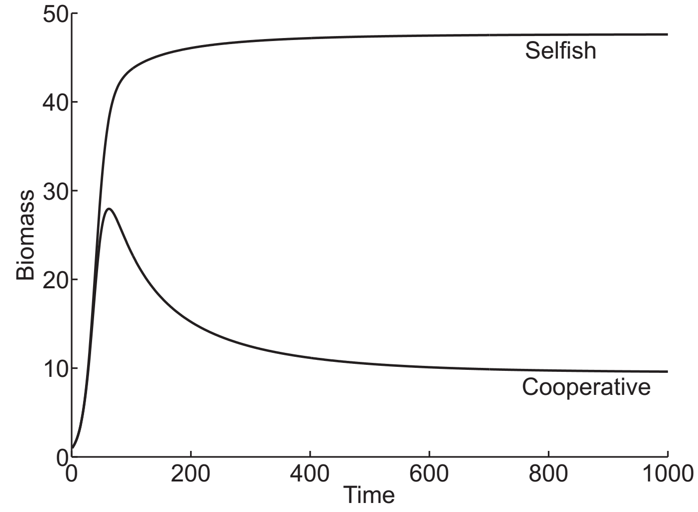
The second, and a key point for our claim, is that the effect of group selection rapidly tends to zero as initial group size increases. Specifically, above a group size of 10, there is no measurable effect at all. Such a result may therefore make the idea of group selection acting on randomly formed groups seem rather implausible as a significant evolutionary pathway. However, the above results only consider the competitive exclusion case; in the coexistence case, the results are somewhat different, as shown in the following section.

4.3.2 The Efficacy of Group Selection under Coexistence Dynamics

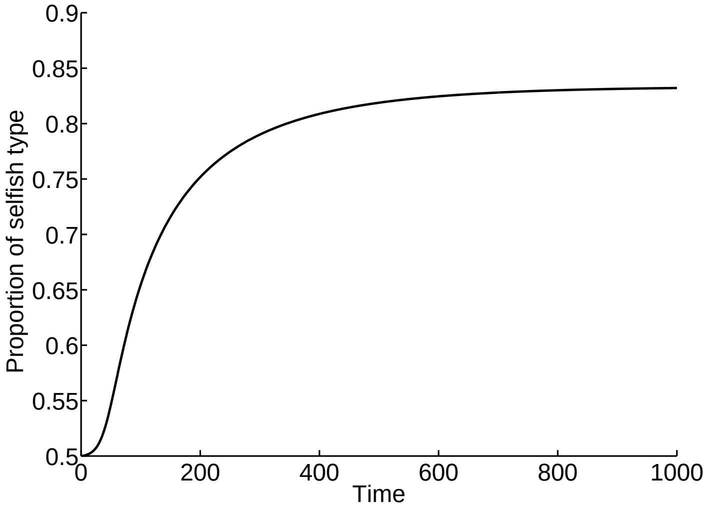
Let us now consider the coexistence dynamics that arise from redefining the selfish strategy as $\alpha_{ss} = 2$ and $\alpha_{cs} = 1.9$, i.e. by swapping the interaction coefficients over from the competitive exclusion case. Figures 4.4**A** and 4.4**B** show how the cooperative type is no longer driven extinct at the within-group equilibrium. In particular, the change in the frequency of the selfish type from an initialisation of 50% shows clear negative frequency-dependent selection resulting in the maintance of cooperation at an above-zero frequency. In other words, the result is a stable coexistence of cooperative and selfish types within a group.

Group selection dynamics under the aggregation and dispersal process are now as shown in the black curve in Figure 4.5. Crucially, in contrast to the competitive exclusion case (shown again in the dotted line), an effect of group selection can be seen over the entire range of group sizes examined. For example, in groups of initial size 50, group selection can be seen to still increase the global frequency of cooperation above the within-group equilibrium. The significance of this observation is that since the within-group equilibrium is the same equilibrium that would be reached in an unstructured population, these results show that group structure is having an effect on population dynamics across a wide range of group sizes.

Finally, to verify that this result is not an artefact of the particular values of α_{ss} and α_{cs} used, the same curves were plotted for a variety of other parameters. Figure 4.6 provides an example of this, where $\alpha_{ss} = 1.99$ and $\alpha_{cs} = 2$ in the competitive exclusion case, vice versa for the coexistence case. These parameters were chosen since they represent stronger within-group selection towards selfish behaviour in the coexistence case than in the previous example. Specifically, the within-group equilibrium frequency of the selfish type in the coexistence case is 98.04%, compared to 83.3% previously. The



(**A**)



(**B**)

Figure 4.4: Coexistence within-group dynamics. A) Biomass of each type. B) Proportion of selfish type.

results in Figure 4.6 show that while an effect of group selection is still seen over a larger range of group sizes in the coexistence case, the magnitude of the effect is reduced compared to Figure 4.5. The reason for this is that variance in group composition is proportional to the frequency of cooperators in the migrant pool in this case, and hence to the corresponding within-group equilibrium frequency, as discussed in detail in the following section.

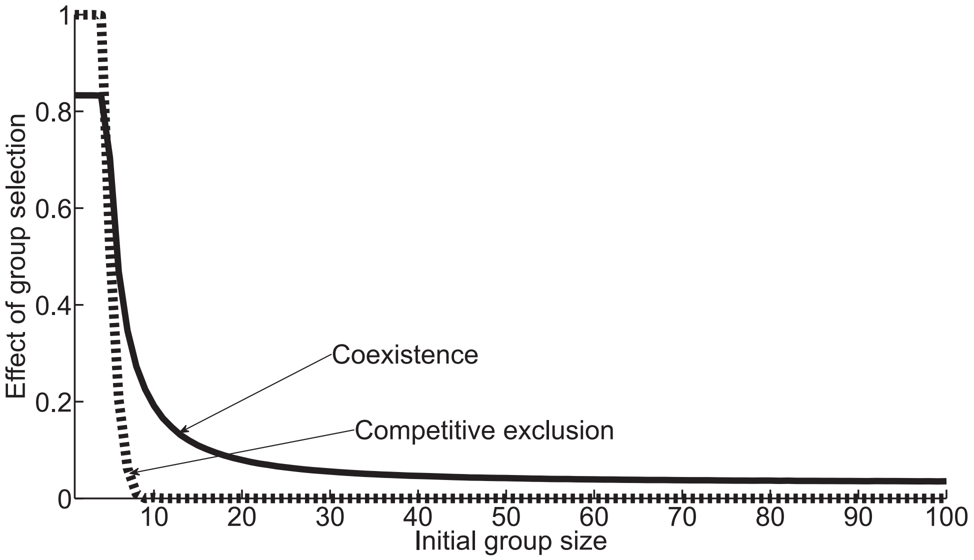


Figure 4.5: Comparing the range of group sizes over which an 'effect of group selection' (see text) can be seen between coexistence and competitive exclusion dynamics.

4.4 Discussion

The results in the previous section demonstrate that where a stable coexistence of types occurs at within-group equilibrium, an effect of group selection on global frequencies can be seen over a much larger range of initial group sizes than in the competitive exclusion case. In particular, as group size increases in the competitive exclusion case, any measurable effect of group selection on the global frequency of cooperation rapidly tends to zero. By contrast, in the coexistence case, some effect on global frequencies is seen over the entire range of group sizes examined. It must be stressed that no particular claim about the magnitude of the effect for large group sizes is made here. Rather, the model implies that there is *some* measurable effect on frequencies over a large range of group sizes; how large this effect may be will depend on the properties of the natural system under consideration. However, the fact that any effect of group selection still exists over a large range of parameters is significant, since it suggests that where within-group dynamics in nature are of the coexistence type, some effect of a group population structure may always be acting.

Coexistence dynamics allow group selection effects to be sustained over a larger range of group sizes because of the effect of migrant pool frequencies on between-group variance. In particular, because group formation constitutes random sampling from the migrant

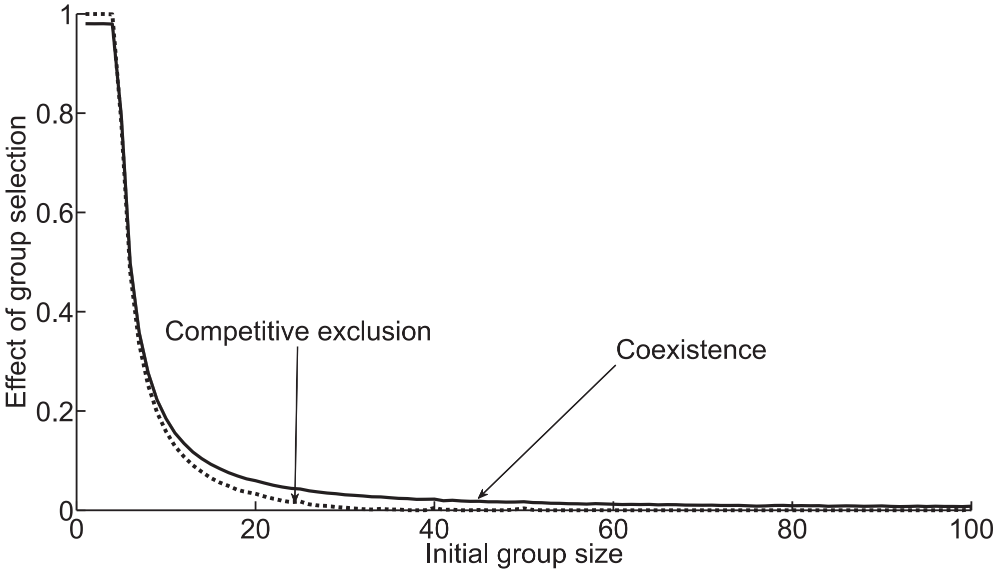


Figure 4.6: Demonstrating that the same qualitative trends arise where within-group selection towards selfish behaviour is stronger in the coexistence case than in Figure 4.5. Here, $\alpha_{cs} = 1.99$ and $\alpha_{ss} = 2$ in the coexistence case, vice versa for competitive exclusion; all other parameters as in Table 4.1.

pool, initial between-group variance can be approximated by the binomial distribution, and is then given by $p_c p_s/S$, where p_c is the proportion of the cooperative type in the migrant pool, p_s the proportion of the selfish type, and S the initial group size (Wilson, 1980, p. 27). Since $p_c + p_s = 1$, it follows that between-group variance is proportional to the frequency of the least frequent type, i.e. variance is maximal when both types are of equal frequency, and zero when one type is at fixation. Therefore, where one type reaches global fixation then there can be no variance and hence no group selection. However, in the coexistence case, where one type cannot reach fixation, it follows that there must always be some variance and hence some possible effect of group selection. The fact that the variance is proportional to the frequency of the least frequent type is illustrated by the difference between Figures 4.5 and 4.6, where the lower within-group equilibrium frequency of cooperators in Figure 4.6 results in a reduced effect of group selection for large group sizes.

A further observation from Figure 4.5 is that a gradient towards an increased effect of group selection also exists over a larger range of group sizes in the coexistence case. Specifically, decreasing group size by a small amount yields an increase in the effect of group selection for groups of size 20 in the coexistence case. However, there is no gradient at this size in the competitive exclusion case. The significance of this is that increasing the effect of group selection increases average absolute individual fitness in the population, due to an increased global level of cooperation. If group size can be partly determined by individual traits then this may provide an adaptive gradient towards smaller groups, increased levels of cooperation, and greater fitness. In the competitive exclusion case, however, such a gradient only exists over a much smaller range of group sizes. The results presented in this chapter therefore suggest that the concurrent evolution of group size and cooperation is more plausible in cases where a stable coexistence of types within groups exists. We investigate this hypothesis in detail in Chapter 6.

4.5 Conclusions

Any group selection process requires there to be a variation in group composition. In aggregation and dispersal style models, this variation arises through the random assignment of individuals from the migrant pool into groups. Consequently, it is often suggested that an effect of group selection on the global frequency of types will only be seen for very small initial group sizes. However, the models on which this claim is based typically only consider within-group dynamics that lead to the competitive exclusion of a cooperative type by its selfish counterpart.

We have shown here that where a stable coexistence of cooperative and selfish types occurs within groups, the range of initial group sizes over which an effect of group selection can be seen is much larger. Whether coexistence or competitive exclusion occurs in a particular social group needs to be checked on a case by case basis. Examples of where this has been done empirically, and coexistence equilibria were found, include cooperative foraging in spiders (Pruitt and Riechert, 2009) and enzyme production in yeast (Gore et al., 2009), as discussed earlier. The results in this chapter suggest that where population structure and social behaviour evolve concurrently, an individual adaptive gradient towards smaller groups and increased cooperation is likely to exist over a much larger range of group sizes under coexistence dynamics. This is in sharp contrast to the competitive exclusion case, where the effects of group selection rapidly reach zero as initial group size increases, excluding the possibility of such a gradient for a large range of parameters.

It has shown in this chapter that the conventional conclusion that group selection effects can only be seen for very small groups rests on the assumption that within-group dynamics lead to competitive exclusion. If instead a within-group coexistence of competing types occurs, then the range of group sizes over which an effect can be seen is much larger. This result follows from the fact that the variance in group composition upon which group selection acts is dependent not only on group size but also on the frequencies of types in the migrant pool. In particular, since neither type can be driven extinct under coexistence dynamics, there will always be some variance in group composition when sampling from the migrant pool, which can then be acted on by group selection. Thus, since it is not necessary to assume that within-group dynamics lead to competitive exclusion, this result shows that group selection can operate in a wider range of conditions than previously realised.

Chapter 5

Evolutionary Game Theory and Multi-level Selection

Evolutionary game theory (Maynard Smith and Price, 1973; Axelrod and Hamilton, 1981; Maynard Smith, 1982) provides a simple and powerful way of conceptualising social evolution, by abstracting away from the details of the underlying genetics of behaviour, and instead focussing on the behavioural phenotypes. In its simplest form, the fitness consequences of a single social interaction between 2 individuals are represented in a payoff-matrix that gives, for a particular strategy, the fitness that will result depending on the partner's strategy. The change in frequency of a behaviour due to selection can then be calculated using the standard replicator equation (Taylor and Jonker, 1978; Hofbauer and Sigmund, 2003):

$$\dot{x_i} = x_i \left((Ax)_i - x^T Ax \right) , \tag{5.1}$$

where x is a vector of strategy frequencies in the population, x_i is the frequency of strategy i, $\dot{x_i}$ is the differential of this, and A is the payoff matrix. A stability analysis of this equation, appropriately parametrised, is then performed to find the equilibrium strategy frequencies in the whole population; this gives the Evolutionarily Stable State (ESS)[1]. Classical evolutionary game theory makes a critical assumption, which is that the population is unstructured. Thus, it is assumed that any individual has an equal probability of interacting with any other individual. At first sight, then, the application of concepts from evolutionary game theory may not seem applicable to group-structured populations. However, this is not the case, because in nearly all models of group-structured evolution it is assumed that interactions within the group are freely-mixed,

---

[1] ESS can also mean Evolutionarily Stable *Strategy*: a strategy which if adopted by all members of the population, cannot be invaded by a rare mutant (Maynard Smith, 1982). Here, we use ESS to mean an equilibrium state of the population that is stable under the dynamics of natural selection.

and so any two *group members* are equally likely to interact. Indeed, a group is often defined as the smallest subset of the population in which interactions are freely-mixed, as is the case with the concept of a trait-group (Wilson, 1975a, 1980). Social interactions within groups can therefore be conceptualised using n-player evolutionary game theory, where n is the group size. Between-group selection, i.e., selection between sets of n-players, can then alter the global ESS. Thus, the ESS in a structured population can be very different from that which solution of the replicator equation yields under freely-mixed conditions.

Many different evolutionary games have been analysed in the literature, corresponding to different kinds of social interaction. The type of game can be categorised in the following manner, according to whether there are one or more ESSs, and whether these ESSs contain only one strategy (pure), or more than one (mixed).

- One pure ESS.
- One mixed ESS.
- More than one pure ESS.
- More than one mixed ESS.

Most models of evolution in group-structured populations implicitly assume that the game under freely-mixed conditions has one pure ESS, corresponding to a selfish type at fixation; the standard equations for the fitness of altruistic and selfish types within groups (e.g., Wright 1945; Levin and Kilmer 1974; Wilson 1975a) can be conceptualised as such a game (Fletcher and Zwick, 2007). Group selection then has the uphill task of supporting cooperation globally when it is not an ESS locally. However, if there is more than one ESS under freely-mixed conditions, representing different outcomes in terms of group fitness, then group selection can perform *equilibrium selection* (Harsanyi and Selten, 1988) between these groups, which is an easier task since between-group variance is more easily maintained (Boyd and Richerson, 1990). We will thus show in this chapter that as we move down the above list, the types of game represent social interactions that are more amenable to group selection, since they allow more between-group variance to be maintained. In line with some other recent work (Bergstrom, 2002; Skyrms, 2004; Doebeli and Hauert, 2005), we will argue that social interactions conventionally modelled as single pure ESS games would be more appropriately modelled by the other classes of game. Where this is the case, we should expect group selection to be more effective than the classical models based on single pure ESS games have suggested. Finally, we will argue that the difference between games with one versus multiple ESSs corresponds, to some extent, to the difference between the fraternal and egalitarian major transitions, *sensu* Queller (1997). Thus, we will suggest roles for different types of game in modelling different aspects of the major transitions.

5.1 Single ESS: the Prisoner's Dilemma and Snowdrift games

Table 5.1: Payoff matrix for the Prisoner's Dilemma game.

	Cooperate	**Selfish**
Payoff to Cooperate	$b - c$	$-c$
Payoff to Selfish	b	0

Games with a single freely-mixed ESS have occupied the vast majority of the biological literature on the evolution of cooperation. Moreover, within this class, the Prisoner's Dilemma game has received the vast majority of attention. The payoff-matrix for this is shown in Table 5.1[2]. The single ESS under freely-mixed conditions is pure, corresponding to every member of the population adopting the selfish strategy (traditionally known as the defect strategy in the game theory literature). Much work has considered how various population structures can produce a different ESS, where some degree of cooperation is stable. Cases that have been considered include those where there are repeated interactions between the same pair of individuals, allowing for reciprocity (Trivers, 1971) such as in the famous "Tit-for-Tat" strategy (Axelrod, 1984). Another well studied case is that in which interactions are single-shot, but occur in a viscous population where relatives interact; this allows kin selection to promote cooperation (Nowak and May, 1992, 1993). Moreover, the Prisoner's Dilemma forms the basis of most multi-level selection models (Fletcher and Zwick, 2007), since it is played locally within groups, but the all selfish ESS can be avoided by between-group selection. This is exactly the situation that the fitness functions in most multi-level selection models capture (Fletcher and Zwick, 2007). Specifically, they consider the n-player Prisoner's Dilemma, where n is the group size; this is also known as the linear public goods game (Bergstrom, 2002). We can generate fitness functions for the n-player Prisoner's Dilemma by multiplying the payoffs in Table 5.1 by the frequency of the corresponding strategy within the group. This yields the following fitness functions (see also Fletcher and Zwick 2007):

$$w_c = w_0 + pb - c \qquad\qquad w_s = w_0 + pb, \tag{5.2}$$

where w_0 is a baseline fitness, p is the proportion of cooperators within the group, b is the benefit of cooperating, and c the cost. It is simple to see from these equations that selfish individuals will always be fitter than cooperators within their same group, because they receive the same benefits pb, but do not pay the cost c. This fact explains why the within-group ESS is all selfish. On the other hand, the greater the proportion of cooperators, the greater the mean fitness (since $b > c$). Consequently, within-group selection *always* favours selfishness, and between-group selection always favours cooperation, exactly as in the vast majority of multi-level selection models.

---

[2] As given in Doebeli and Hauert (2005)

In the Prisoner's Dilemma, the two levels of selection are therefore exactly in opposition, which provides a worst-case scenario for any effect of group selection to be seen. However, even within the category of single ESS games, this does not have to be the case. For example, in the Snowdrift game (Sugden, 1986; Doebeli and Hauert, 2005), or Hawk Dove (Maynard Smith, 1982) game, negative frequency-dependent selection operates, such that some level of cooperation is supported within a group at equilibrium. The pay-off matrix for this game is shown in Table 5.2. The story behind this game is that two car drivers are trapped in a snowdrift, and can either get out and shovel (cooperate) or stay in their car (selfish). The difference from the Prisoner's Dilemma is that if the other driver does not cooperate, then it is better for you to shovel, for at least then you will get home, even though less effort would be required if both cooperatively shovelled (Sugden, 1986). On the other hand, if the other driver shovels, it is individually advantageous for you to not do so but to cheat and remain in the warmth of your car. The equilibrium in a population of players is therefore some individuals cooperating, and some cheating, in contrast to the all selfish equilibrium of the Prisoner's Dilemma. The Snowdrift game can therefore be seen as representing a relaxed version of the Prisoner's Dilemma, in which unilateral cooperation is fitter than mutual defection, but unilateral selfishness is still individually fitter than mutual cooperation (Doebeli and Hauert, 2005). Indeed, the pay-off matrix in (Doebeli and Hauert, 2005) can represent both types of game; if $1/2 < b < c$ it yields the Prisoner's Dilemma (Doebeli and Hauert, 2005).

Table 5.2: Payoff matrix for the Snowdrift game (Doebeli and Hauert, 2005).

	Cooperate	**Selfish**
Payoff to Cooperate	$b - c/2$	$b - c$
Payoff to Selfish	b	0

What kinds of biological interaction might give rise to the Snowdrift game, rather than the Prisoner's Dilemma? Essentially, it those where there is some direct benefit to the cooperator of her actions, that is not shared with other group members. When considering public goods, this would correspond to a producer being able to keep some fraction of the good that she produces for herself, with the rest being shared with the group. When this happens, cooperators receive a greater *per capita* share of the benefits than selfish individuals and so are, all other things being equal, fitter. However, for negative frequency-dependent selection to operate, it must also be the case that the fitness of cooperators declines as they increase in frequency. This can happen if the costs of cooperating are fixed, but the benefits become discounted with increasing levels of cooperation (Hauert et al., 2006; Gore et al., 2009). Thus, although at low frequencies the benefits of cooperation may outweigh the costs, at higher frequencies these benefits become discounted such that this is no longer the case. The result is a single mixed ESS under freely-mixed conditions, which gives a polymorphism of cooperative and selfish behaviours within the group.

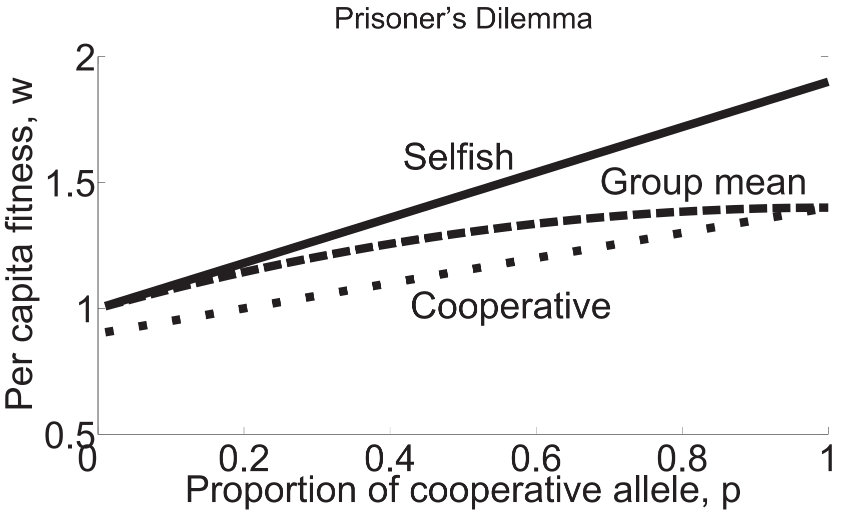
Biological interactions for which the Snowdrift game may be more appropriate than the Prisoner's Dilemma can be found across taxa and include, as examples, enzyme secretion in yeast (Gore et al., 2009) and viruses, antibiotic resistance in bacteria (Dugatkin et al., 2003, 2005), social foraging in spider colonies (Pruitt and Riechert, 2009), and sentinel behaviour in mammals (see also Doebeli and Hauert 2005 for a review) . In general, the Snowdrift game applies to any situation where some level of cooperation is individually advantageous, but where the individual benefit tails off with increasing frequency. It should be noted that non-linear benefits to cooperation lead to negative frequency-dependent selection even in the absence of population structure or a varying global carrying capacity, which were the mechanisms considered in Ross-Gillespie *et al.*'s (2007) recent treatment, and hence negative frequency-dependent selection can occur locally within single groups and not just at the metapopulation level as has been considered in previous works.

Crucially, a conflict between levels of selection still exists in the Snowdrift game, since group productivity would still increase if more group members were cooperative than is supported under within-group selection, a point that has very rarely been noted in the literature. By failing to take account of this fact, the effect of Snowdrift style interactions on the efficacy of group selection has been overlooked (see Dugatkin et al. 2005 for an exception where Snowdrift style interactions are considered in a group structured population). In the next section, we will explicitly show that while some degree of cooperation can be maintained in the Snowdrift game by within-group selection, between-group selection can still increase this further, and will still be opposed by individual selection within groups. We then show that Snowdrift style interactions increase the range of conditions over which an effect of group selection can be seen, due to the polymorphic within-group equilibrium allowing some between-group variance to be regenerated. This is an instance of exactly the same mechanism that we considered with respect to coexistence of species in the Lotka-Volterra model of Chapter 4.

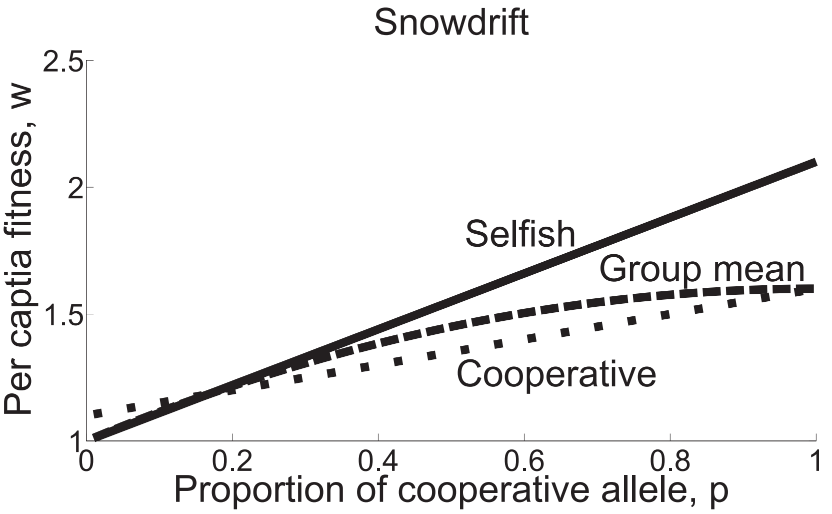
The fitness functions for the Snowdrift game within groups, derived from Table 5.2, are as follows:

$$w_c = p\left(b - \frac{c}{2}\right) + (1-p)(b-c) + w_0 \qquad\qquad w_s = pb + w_0, \tag{5.3}$$

where w_c is the fitness of group members with the cooperative allele, w_s the fitness of those with the selfish allele, p the proportion of the cooperative allele within the group, b the benefit of cooperation, c the cost, and w_0 a baseline fitness in the absence of social interactions. In these equations $b/c > 1$ produces the Snowdrift game, and $0.5 < b/c < 1$ the Prisoner's Dilemma (Doebeli and Hauert, 2005). These fitness functions are shown graphically in Figure 5.1, which highlights the fact that in both games, the mean fitness of group members would be higher with a greater proportion of cooperators than is supported at (within-group) equilibrium. Thus, a conflict between individual- and group-level selection exists in both games.



(**A**)



(**B**)

Figure 5.1: *Per capita* fitness as a function of the frequency of the cooperative allele within the group in *a*) Prisoner's Dilemma, *b*) Snowdrift, games. A polymorphic equilibrium exists in the Snowdrift, but not Prisoner's Dilemma, game. A polymorphic equilibrium is indicated graphically where the lines of the fitness functions cross, meaning that both types have equal fitness. The Prisoner's Dilemma was produced by setting $b = 0.9$ and $c = 1$ in Equation 5.3. The Snowdrift game was produced by setting $b = 1.1$ and $c = 1$.

5.2 Prisoner's Dilemma and Snowdrift games in group-structured populations

Most previous work on the Snowdrift game has considered its operation in unstructured populations (but see Hauert and Doebeli 2004 and Számadó et al. 2008, discussed below). Here, we consider its operation in multi-generational, competing, groups of size n.

We use the aggregation and dispersal model of the previous chapter, where groups of initial density n_0 are formed by random sampling without replacement from a migrant pool of total density 1000. Reproduction then occurs within these groups for a number of generations, with fitnesses as given by Equation 5.3. The density of a type within the group in the next generation depends on its fitness and is given by Equation 5.4:

$$n_c(g+1) = n_c(g)w_c \qquad\qquad n_s(g+1) = n_s(g)w_s, \tag{5.4}$$

where g is the generation within the group, n_c is the density of the cooperative type, and n_s the density of the selfish type within the group. We allow these densities to be continuous numbers, as they are in the Lotka-Volterra equations. After t generations the groups disperse and form a new migrant pool. We cap the size of the migrant pool to be equal to the initial population size, thereby allowing for a global carrying capacity. This is done by rescaling the migrant pool back to the initial population density of 1000, keeping the frequency of cooperative and selfish types the same. At this stage, we round the densities of the cooperative and selfish types to integers in order to return the population to its original integer density. Thus, our model is not individual-based, but instead uses growth equations on continuous densities, in a manner analogous to that in Fletcher and Zwick (2007).

5.2.1 Results under random group formation

Figure 5.2 shows the effect of group selection on the global frequency of cooperation in the Snowdrift and Prisoner's Dilemma games, under randomly formed groups of varying initial sizes. The simulations were run for 1000 aggregation and dispersal cycles, which was a sufficient length of time for a global equilibrium to be reached, and were repeated 30 times. We define "effect of group selection" as the (absolute) difference between the equilibrium frequency of cooperation in a freely-mixed population, and the global frequency of cooperation in the group metapopulation after 1000 aggregation and dispersal cycles (we take the average global frequency over the last 100 aggregation and dispersal cycles, since the allele frequencies slightly fluctuate between cycles due to the random group formation process). In the Snowdrift game, the greatest possible effect of group selection decreases as the benefit-to-cost ratio increases. For example, when $b : c = 1.1$ the equilibrium frequency of cooperation in a freely-mixed population is around 19%, but when $b : c = 1$ it is around 66%. Thus, the absolute effect that group selection can have is lower in this second case, according to our metric. This is why the curves for higher $b : c$ ratios start off lower on the y-axis in Figure 5.2.

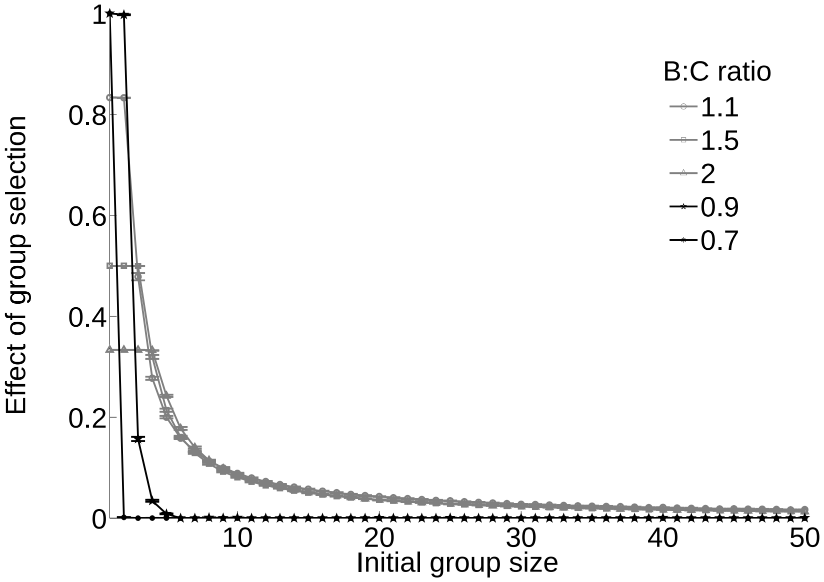


Figure 5.2: Effect of initial group size in the Snowdrift and Prisoner's Dilemma games under random group formation, using Equation 5.3. The grey lines show the Snowdrift game with c fixed at 1 and various values of b. The black lines show the Prisoner's Dilemma produced with $b = 0.9$ (upper line), and $b = 0.75$ (lower line). Error bars show the standard error over 30 trials.

We find that under the Prisoner's Dilemma game, there is no effect of group selection on cooperation over a large range of initial group sizes. By contrast as initial group size decreased in the Snowdrift game, some global increase in cooperation could be seen over the entire range of group sizes examined. The distinction between pure and mixed ESS games (Prisoner's Dilemma and Snowdrift, respectively) is exactly the same as that between competition and coexistence in the Lotka-Volterra equations, as examined in Chapter 4. Specifically, a mixed ESS under freely-mixed conditions, and hence within a group, maintains the possibility of between-group variance being generated when the groups are reformed from the migrant pool. This can occur even when the group composition is close to the ESS, as happens when the initial group size is large. Consequently, a small decrease in initial group size can increase this variance over a large range, and hence provide a population structure that selects for greater cooperation. By contrast, in games with a single pure ESS, the possibility for any between-group variance is destroyed as the groups approach that ESS. Our model results illustrate that this qualitative difference between the Snowdrift and Prisoner's Dilemma games holds for a range of benefit-to-cost ratios in both games.

Our results on the effect of group-structure in the Snowdrift game should be compared with those of Hauert and Doebeli (2004), who considered the game played in a viscous population occupying a spatial lattice. They claimed that spatial structure decreases the frequency of cooperation in the Snowdrift game, compared to the well-mixed case. Thus, from their results we might predict that group structure would likewise decrease the equilibrium frequency of cooperation. Our results have shown this prediction to be incorrect, however, since we have found that group structure *does* increase the global level of cooperation in the Snowdrift game. A recent paper by Számadó et al. (2008) can explain this apparent contradiction. Specifically, Számadó et al. show that playing a game on a spatial lattice has two effects: it introduces correlations in the strategy of interacting individuals, and it reduces the neighbourhood in which competitive interactions take place[3]. They found that the introduction of correlations in the strategy of players (that is, between-group variance or relatedness greater than zero) always had a positive effect on the frequency of cooperation, and that it was the limited neighbourhood of interactions that hindered cooperation in Hauert & Doebeli's model. Limited neighbourhood has analogous effects to the number of generations spent within groups before dispersal in our model. Specifically, a small neighbourhood (coupled with a population at local carrying capacity) means that *competitive* interactions take place on a small scale, likewise for a large number of generations before dispersal in our model, as discussed in Chapter 3. It is now well-known that a limited range of competitive interactions hinders the evolution of cooperation (Wilson et al. 1992; West et al. 2002, again as discussed in Chapter 3), and this applies in both the Snowdrift and Prisoner's Dilemma games (Számadó et al., 2008). The global level of cooperation increased in our model in the Snowdrift game, compared to the well-mixed case, because the group structure introduced correlations in the behaviour of interacting individuals (which always favours cooperation, *contra* to the suggestion of Hauert and Doebeli 2004) whilst regular dispersal prevents the effects of limited neighbourhood.

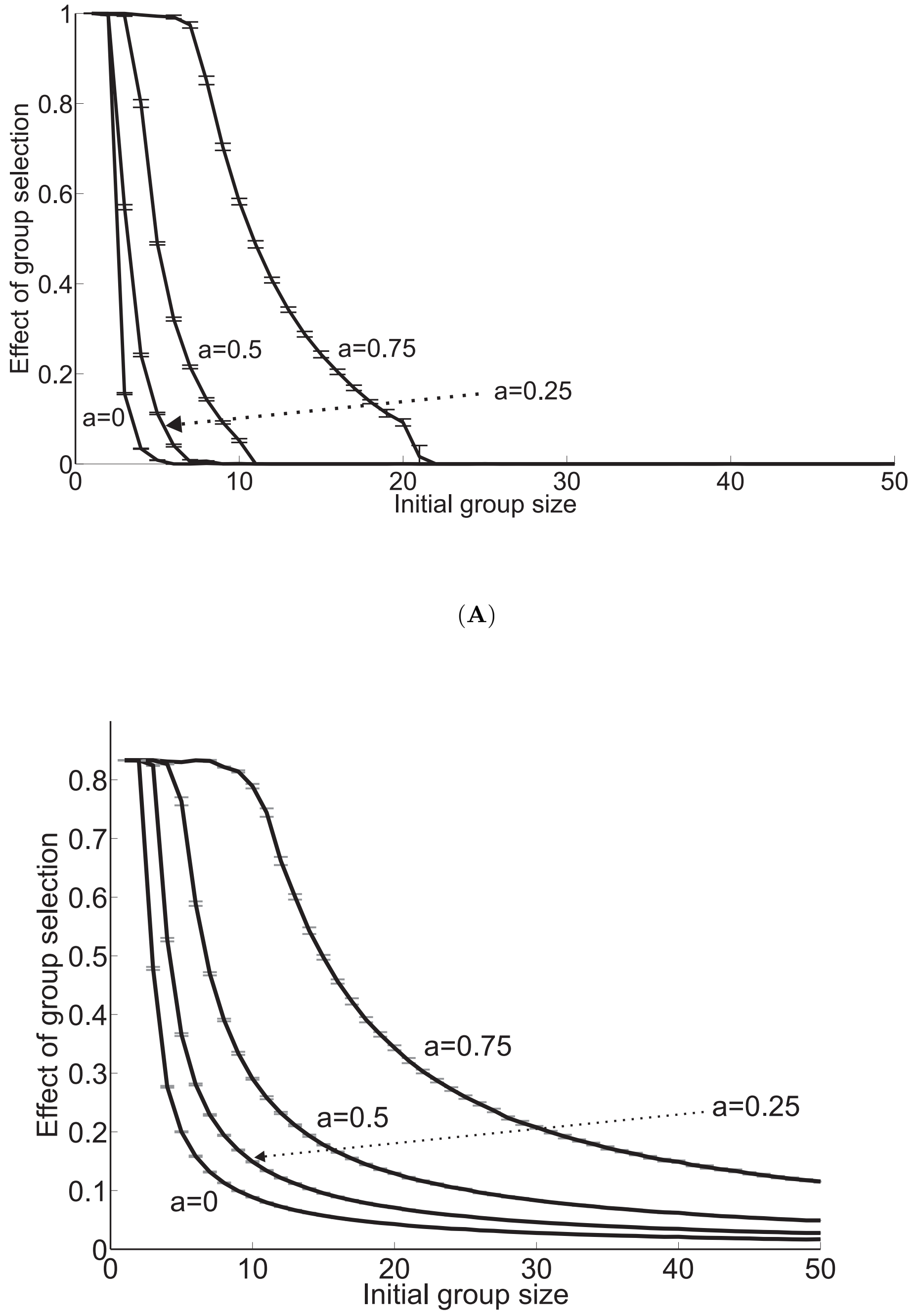
5.2.2 Results under assortative group formation

In the previous section we have considered groups that are formed randomly. However, it is also possible for groups to be formed assortatively, such that individuals with the same behavioural phenotype tend to found groups together. Such assortativity arises if the groups are founded by kin, such as sibling groups. It can also potentially arise if the cooperative allele has some pleiotropic effect on the habitat preference of its bearers (Hamilton, 1975). We have investigated the effects of assortativity to see if the distinction between the Snowdrift and Prisoner's Dilemma games discussed above can still be seen. To do so, we use the model described above, but introduce an assortativity index,

---

[3] In the lattice style of spatial model considered by Hauert & Doebeli, the fitness of a focal individual is compared with that of one of k adjacent neighbours on the lattice. An offspring of the selected neighbour takes over the focal individual's site with a probability proportional to the difference in fitness (payoff) between the focal individual and that neighbour. See Számadó et al. (2008).

a, as follows. The first founder of a group is chosen randomly, as before. Thereafter, the next founder of a group is chosen to be of the same behavioural type as the previous member added with probability a, otherwise the next founder is chosen randomly. Thus, when $a = 1$ all founders of the group will be of the same behavioural type, whereas when $a = 0$ all founders of the group will be chosen randomly as in the previous section.



(**B**)

Figure 5.3: Effect of varying degrees of assortative group formation in the Prisoner's Dilemma (A), and the Snowdrift game (B).

Regardless of the degree of assortativity in group formation, we can see from Figure 5.3 that the same qualitative distinction between the two games holds. That is, any effect of group selection still rapidly tails off to zero as the initial group size increases in the Prisoner's Dilemma, regardless of assortitivty. By contrast, in the Snowdrift game the decline in between-group variance was much more gentle, and some effect of group selection could be seen over the entire range of initial group sizes we examined. Thus, the Snowdrift game preserving between-group variance applies equally under population structures where group formation is random or assortative. We also find that as the assortativity increases, the effect of group selection tails off less rapidly with increasing initial group size. This is to be expected from classical theory, since assortativity increases the initial between-group variance, and so counteracts the effect of increasing initial group size.

5.3 A classification of games and their effect on group selection

The type of social interaction can be classified in game theory along the following two axes:

- whether there is more than 1 ESS;
- whether the ESS is pure or mixed.

In general, much of the biological literature on social evolution has been preoccupied with the case where there is one pure ESS, corresponding to the most selfish type at fixation and mean population fitness at its lowest. Such a case corresponds to both strong altruism, and to the classical conception of tragedy of the commons / public goods scenarios. Indeed, explaining strong altruism has often been *the* motivation for investigating the possibility of group selection, even though there are many other kinds of traits that could be promoted by between-group selection (Wilson, 1997). As has been shown in both this and the previous chapter, the case of one pure ESS corresponds to something of a worst-case for group selection, not only because it represents strong within-group selection, but because it necessarily curtails the range of parameters over which between-group variance can be generated compared to other games, a point that we have developed in this dissertation for the first time. Furthermore, it represents an extreme kind of biological cooperation, where the social dilemma between the "good of the individual" and the "good of the group" is most acute (Doebeli and Hauert, 2005). In particular, it assumes that there is not even a small amount of unilateral benefit to performing a cooperative act, and that even if cooperation is fixed in a population, it must still be vulnerable to a cheating type arising by mutation. These two worst-case assumptions are relaxed in the Snowdrift, and Stag Hunt games, respectively.

Table 5.3: Payoff matrix showing relationship between the Prisoner's Dilemma, Snowdrift, and Stag Hunt games. R is the reward for mutual cooperation, P the punishment for mutual defection, S the sucker's payoff for cooperating when your partner does not, and T the temptation for defecting when your partner cooperates. $T > R > P > S$ defines the Prisoner's Dilemma, $T > R > S > P$ the Snowdrift game, and $R > T > P > S$ the Stag Hunt (see, for example, Santos et al. 2006.)

	Cooperate	**Selfish**
Payoff to Cooperate	R	S
Payoff to Selfish	T	P

The payoff matrix in Table 5.3 shows the relationship between the Prisoner's Dilemma, Snowdrift, and Stag Hunt games (Santos et al., 2006). The Snowdrift game relaxes the assumption that there are no unilateral benefits to performing a cooperative act. This is shown in the payoff matrix by the fact if the actor's partner is selfish, then the actor will be fitter by being cooperative, rather than selfish himself ($S > P$). As discussed above, this leads to a single, *mixed*, ESS in a freely-mixed population. Such a scenario can occur when cooperators can internalise some of the benefit that they produce (Gore et al., 2009), for example, keep some of the good that they produce for themselves rather than sharing it with the whole group. Where such a mixed ESS is supported under within-group selection, we have argued that between-group selection will have some effect of globally raising the frequency of cooperation over a larger range of conditions. Thus, even loose forms of population structure can affect the equilibrium frequency of cooperation in mixed ESS games.

A numerical example payoff matrix for the Stag Hunt game is shown in Table 5.4; this is a particular class of coordination game where cooperation involves coordination of behaviour. The intuitive story behind the game is that an individual can choose to hunt either a stag or a hare (Skyrms, 2004). Successfully hunting the stag yields the highest fitness, but requires that both individuals cooperate to do so; if one individual does not, then the other would have been better off (selfishly) hunting the hare. The Stag Hunt game relaxes the assumption that if all other individuals cooperate, then it pays to behave selfishly. This is shown in Table 5.3 by the fact that if the actor's partner cooperates, then the actor is fitter by himself cooperating rather than changing his behaviour to selfishness ($R > T$). In other words, the temptation to defect from mutual cooperation is removed (Santos et al., 2006). This results in all cooperate being a stable equilibrium as well as all selfish, with all cooperate leading to greater fitness for all of the individuals. There are thus *two* pure ESSs of different total fitness in a freely-mixed population, both hunt hare, or both hunt stag. If unilateral selfishness is fitter than unilateral cooperation ($T > S$), then the all selfish ESS will have the largest basin of attraction, and so will be reached from a larger range of starting behavioural frequencies in the population. In the language of classical game theory, all cooperate is the payoff dominant equilibrium (Harsanyi and Selten, 1988), since it yields the highest

Table 5.4: Payoff matrix for the Stag Hunt game, adapted from Skyrms (2004). There are two equilibria: both hunt stag (optimal, payoff dominant), or both hunt hare (suboptimal, risk dominant)

	Stag	**Hare**
Payoff to Stag	5	1
Payoff to Hare	3	2

payoff (fitness) to each individual. On the other hand, all selfish is the risk dominant equilibrium (Harsanyi and Selten, 1988). This is intuitively because it is safer to play selfish if you do not know the behaviour of your partner (the formal definition of risk dominance is that it has the largest basin of attraction Harsanyi and Selten, 1988).

The choice of risk- or payoff- dominated equilibrium by the individuals is known as equilibrium selection (Harsanyi and Selten, 1988). Group structure can increase the likelihood that the equilibrium optimal for the group (e.g., stag) is reached, by the following mechanism. Groups whose members cooperatively hunt stags will outcompete those whose uncoordinated members selfishly (or unilaterally) hunt hares (Canals and Vega-Redondo, 1998; Weibull and Salomonsson, 2006). This means that the stag strategy will increase in frequency in the global population when the groups mix. As a result, when new groups are formed they will, on average, contain a higher proportion of stag hunters, and of more of the groups will be in the basin of attraction for stag. Consequently, the greater mean fitness of the stag ESS can set up a positive feedback process that leads to its fixation in all groups. In this way, group selection selects for the payoff dominated equilibrium (Canals and Vega-Redondo, 1998). From a biological viewpoint, this is intuitive; group selection is favouring those groups whose members behaviour is well coordinated, and workers such as Wilson (1997) and Seeley (1997) have argued that behaviours such as coordination are a group-level adaptation (but see also Chapter 2 of this dissertation). Moreover, this kind of group selection is likely to work over a broader range of conditions than is conventionally understood in biology, since the initial between-group variance when the groups are founded can be magnified by within-group selection taking the groups to different ESSs (Boyd and Richerson, 1990; Wilson, 1992). By contrast, in a single ESS game, within-group selection will drive all groups towards the same ESS, thereby reducing any initial between-group variance. Thus, we would expect games with more than one ESS to, all other factors being equal, lead to the generation of more between-group variance than single ESS games. Therefore, an effect of group selection should be seen over a larger range of conditions.

Interestingly, between-group selection is favoured in multiple ESS games when the groups *do* reach equilibrium, so enough generations must be spent within the groups before dispersal (or equivalently, the intergroup migration rate must be sufficiently low) to allow equilibrium behavioural frequencies to be reached (Boyd and Richerson, 1990). This is the opposite of pure ESS games, where the groups must disperse *before* the equilibrium is reached if an effect of group selection is to be seen. Rapid within-group selection therefore supports between-group selection in multiple ESS games. Boyd and Richerson (1990) have suggested that processes of cultural evolution, in which group social norms are established and rapidly become fixed by imitation, should be particularly conducive to such group selection. Another way in which within-group selection can happen on a much faster time scale than intergroup mixing is if the groups start out as separately isolated populations that slowly spread out into contact with each other. Within-group selection could then reach an ESS before the populations meet, and between-group selection would then favour the population that reached the ESS with greatest mean fitness (Snowdon et al. 2009, see also McElreath et al. 2003). Such a process could occur with genetic evolution, without requiring cultural imitation (Snowdon et al., 2009).

However, although multiple equilibria games such as the Stag Hunt have been well studied in economics (e.g., Harsanyi and Selten 1988; Van Huyck et al. 1990; Kandori et al. 1993; Oechssler 1997), including the effects of group structure on equilibrium selection (Canals and Vega-Redondo, 1998; Weibull and Salomonsson, 2006), little attention has been paid to them as biological models of social behaviour (for exceptions see Weibull and Salomonsson 2006 and Pacheco et al. 2009). We, along with authors such as Skyrms (2004) and Santos et al. (2006), argue that this is a mistake, and that coordination of behaviour is a fundamental kind of biological social interaction whose evolution does indeed need explaining. One important example of the general coordination game is establishing word-meaning pairs in the origin of human language (where the origin of language is itself a major transition in evolution; Maynard Smith and Szathmáry 1995) (Szamado and Szathmary, 2006). Pacheco et al. (2009) suggest that the stag-hunt game is a model of the coordinated hunting actions of a range of animals, including lions and chimpanzees. It has also been suggested as a model of symbiotic interactions, where the relationship between symbiont and host can evolve into one of either mutualism (stag), or parasitism (hare) (Law, 1991; Maynard Smith and Szathmáry, 1995). Population structure plays a key role in such interactions, since localised social interactions coupled with global competition (for example, an aggregation and dispersal process) can allow the payoff dominant (cooperative), rather than risk dominant (selfish) equilibrium to be reached, just as it allows cooperation to evolve in single ESS games.

Our argument is not simply that cooperation is easier to evolve in types of game other than the Prisoner's Dilemma. Rather, it is that the range of group structures that have an effect on social evolution is likely to be much larger if interactions within groups take the form of a mixed or multiple ESS game. In particular we suggest that the three types of game discussed above, to wit, one pure ESS (Prisoner's Dilemma), one mixed ESS (Snowdrift), and multiple pure ESSs (stag-hunt), can be viewed on a sliding scale, in that order, of how amenable they are to group selection. The fundamental reason for this is that as this scale progresses, so does the effect of within-group selection on preserving between-group variance. In games with one pure ESS, within-group selection acts to, in the limit, completely destroy any initial between-group variance. In the Snowdrift game, some between-group variance can be preserved when the groups are reformed, since one type cannot be driven extinct. Where there are multiple ESSs, the initial variance in group composition can actually magnified by within-group selection taking different initial group compositions to different equilibria. It has been argued that the Snowdrift game may be a more appropriate model of public goods scenarios than the Prisoner's Dilemma in many cases (Doebeli and Hauert, 2005; Gore et al., 2009), and it has also been argued that the Stag-Hunt is a more appropriate model of interactions such as cooperative hunting Skyrms (2004); Pacheco et al. (2009). These assertions about how we should best conceptualise and model cooperative interactions have the important implication that group selection should be more effective than previous models, based nearly always on the single pure ESS case, have indicated.

5.4 The role of different game types in modelling the major transitions

Following Queller (1997), the major transitions in evolution can be divided into two classes: fraternal and egalitarian. In the fraternal case, the lower-level units are of the *same* type and are hence fungible, that is, they can stand in for each other in terms of function. This creates the opportunity for both benefit and conflict. The potential benefit is that a reproductive division of labour can occur, for each unit carries the same genes (bar mutation), and so not all units need to reproduce to pass these genes on (Queller, 1997). Examples of such reproductive division of labour include the germ-soma separation in multicellular organisms, and the existence of sterile worker castes in eusocial insects (Maynard Smith and Szathmáry, 1995). However, this also creates the opportunity for conflict, for selfish mutants can arise that reproduce at a faster rate, to the detriment of group productivity (Michod, 1999). Indeed, by definition a reproductive division of labour involves *strong* altruism.

For modelling such a division of labour, we suggest that the Prisoner's Dilemma is the most suitable game, for in its n-person variety it represents strong altruism whenever the benefit:cost ratio is less than the group size (Fletcher and Zwick, 2007). It has been shown here, and elsewhere (e.g., Fletcher and Zwick 2007), that the conditions for cooperation to evolve in this game are very stringent, for it represents the strongest form of social dilemma. In particular, group formation must typically be highly assortative[4] if between-group selection is to overcome within-group selection in this case. However, in many of the fraternal transitions, the groups are founded by one or a few individuals. For example, metazoans are cell-groups that develop from a single cell. Development from

---

[4] Assortativity can also build up as reproduction occurs within groups Fletcher and Zwick (2004).

a single cell means that all the variance is between cell-groups, and there is no within-group variance or selection (bar mutations), hence cooperation as defined by a Prisoner's Dilemma game can indeed evolve (Maynard Smith and Szathmáry, 1995; Queller, 1997, 2000). Similarly, eusocial insect colonies develop from one or a few queens, and so there can again be sufficient between-group variance, and low within-group variance, to allow a reproductive division of labour to evolve under a Prisoner's Dilemma model.

Reproductive division of labour, however, is not the only cooperative trait relevant to the fraternal transitions. In particular, "Tragedy of the Commons" type situations, such as public goods production and resource exploitation, can occur. As argued above, in such cases the Snowdrift game may in fact be a more appropriate model of within-group competition, or in general, a game with a mixed ESS. The results presented in this chapter show that group selection is effective over a much wider range of parameters in such cases. This agrees with the fact that such types of cooperation are seen much more frequently in nature than are strongly altruistic traits (such as division of labour). While it can be argued that reproductive division of labour is the hallmark of a fraternal transition, other types of cooperation such as efficient resource utilisation are also essential, and are likely to evolve first.

We have postulated here that the Prisoner's Dilemma and Snowdrift games are suitable models of different types of cooperation that occur during the fraternal transitions. What, then, of the egalitarian transitions? In these transitions (for example, the packaging of genes into chromosomes, and the origin of the eukaryotic cell Queller, 1997), the pre-transitional units are of different types, and perform different functions. Reproductive division of labour is therefore not possible, since an individual of one type cannot pass on the genes of another. However, since the pre-transitional organisms are of different species, they can already perform different functions when brought together, and this combination of functions can provide the initial selective advantage to group membership (Queller, 1997).

We suggest that one of the difficulties is then in finding sets of compatible types that perform complimentary, and hence synergistic, functions. We then argue that such situations are most appropriately modelled by a game that has multiple mixed ESSs. The ESSs should be mixed, since a coexistence of different types occurs by definition in the egalitarian transitions. Each mixed ESS can then represent a different (compatible) combination of types. Some of these ESSs will yield a higher group fitness (in terms of group productivity) than others, and group selection can act between these (Boyd and Richerson, 1990; Wilson, 1992). However, there may be a combinatorial number of such ESSs (Watson et al., 2009; Mills and Watson, 2009), which can frustrate the evolution of high fitness groups through the sheer number of possible combinations. Moreover, some of these ESSs may be suboptimal but risk dominant, causing the high fitness ESSs to have small basins of attraction; this would be a generalisation of the Stag Hunt game to mixed ESSs. Thus, between-group selection can find payoff dominant ESSs of high group fitness. It should be noted there may still be a “Tragedy of the Commons” scenario in such transitions as well. For example, the different species may share a limiting resource that their growth most depends upon; see, for example, Szathmáry 1999 where this occurs due to different types of replicator sharing common metabolic products on a surface or in a protocell. Thus, both single-ESS games representing a “Tragedy of the Commons”, and multi-ESS games representing coordination of function, may be appropriate. The appropriateness, of course, needs to be carefully checked against biological facts in each case. In general, we stress the appropriateness of different kinds of game for modelling different types of social interaction, and think that, like economists, social evolutionists would do well to explicitly consider other types of social interaction apart from the n-player Prisoner’s Dilemma.

5.5 Conclusions

In this chapter, we have taken a game theoretic approach to conceptualising the social interactions that take place within groups. Although game theory is more abstract than the models considered in the previous chapters, we have used this abstraction to shed light on the underlying nature of social interactions. In particular, nearly all previous models of evolution in group structured populations implicitly assume that social interactions within a group take the form of a Prisoner’s Dilemma (Fletcher and Zwick, 2007). Yet, this assumption represents a worst case social dilemma, and we have shown that it necessarily curtails the range of conditions over which between-group variance can be generated. The social dilemma in the Prisoner’s Dilemma game can be relaxed in two ways. First, unilateral cooperation can be fitter than mutual selfishness, as in the Snowdrift game. This creates a single mixed ESS, which allows the possibility of between-group variance being generated when the groups are reformed. Second, mutual cooperation could be fitter for each individual than unilateral selfishness, making all cooperate an ESS, as in the Stag Hunt game. This creates two ESSs, and group selection can perform equilibrium selection between them, which favours the payoff dominant all-cooperate ESS. The existence of more than one ESS provides another mechanism by which between-group variance can be generated.

We would argue that taking a game theory approach highlights these different kinds of social interactions, and that the default prisoner’s Dilemma case may not be appropriate for a range of biological cooperation. For example, if cooperators can internalise some of the benefit in a public goods scenario, then the mixed ESS Snowdrift game may be a better model (Doebeli and Hauert, 2005; Gore et al., 2009). The multiple ESS Stag Hunt game may be a better model of ecological interactions such as cooperative hunting (Pacheco et al., 2009), or symbiotic living (Law, 1991; Maynard Smith and Szathmáry, 1995). Game theory gives us simple equations to model these different kinds of social interactions. By contrast, most work on multi-level selection uses only one kind of fitness function, which dates back to Wright (1945) and implicitly represents the Prisoner's Dilemma. We have argued that other types of social interaction are inherently more conducive to group selection, not because it is easier to evolve cooperation in them, but because they generate more between-group variance. Thus, population structure is more likely to have some effect on the evolution of social traits in such games.

Chapter 6

The Concurrent Evolution of Population Structure and Cooperation

In this chapter, we develop a model that illustrates how social niche constructing individuals can evolve from living in a population structure where selfish behaviour dominates to one where cooperation is selected for. This model builds on the one presented in Chapter 3, by allowing new initial group size preferences to be generated over the course of evolution by mutation. We can therefore use this model to investigate conditions under which changes in population structure can occur by selection for a series of small mutations, in contrast to the large differences in group sizes that we considered in Chapter 3. We thus consider not only the joint equilibrium of structure and social behaviour, but also whether an individual adaptive gradient exists towards that equilibrium. In line with the argument developed in chapters 4 and 5, we hypothesise that such a gradient should occur from a much wider range of initial conditions where within-group selection on social behaviour exhibits negative frequency-dependence. In this chapter, we model the difference between frequency-dependent and directional selection on social behaviour using the Snowdrift and Prisoner's Dilemma games, respectively. We would thus predict that the evolution of initial group size towards that supporting greater cooperation should be more plausible where within-group social interactions take the form of the Snowdrift game.

6.1 Introduction

Typically, theoretical studies of social evolution selection have considered whether a given population structure can provide sufficient between-group variance for an individually-costly, but group-beneficial, cooperative trait to evolve (or equivalently, whether there is a sufficient degree of relatedness between interacting individuals). Thus, the population structure, and hence the amount of between-group variance that can arise, is assumed to be static, with only social behaviour being subject to evolutionary modification. By contrast, we are interested in the following question: how do the population structures that support cooperation themselves arise? One possibility is that they are provided by physical features of the environment. Indeed, organisms very rarely live in perfectly well-mixed conditions, and so some kind of population structure providing localised interactions will typically be present (see also the concept of Wilson's trait groups (Wilson, 1975a, 1980)). For example, a resource may be distributed in a patchy manner across a landscape, creating groups of organisms that feed from the same patch. Cooperative feeding traits, such as a reduced consumption rate to increase yield (Pfeiffer et al., 2001; Kreft, 2004) may then evolve through between-group variance and competition, as discussed above. However, in addition to that provided by the environment, we explicitly account in this model for the fact that individuals can also evolve their own population structure. In particular, we consider the evolution of a genetic trait that affects between-group variance and hence the balance between levels of selection. An example of such a trait includes a preference for the number of other conspecifics with which an individual founds a group, for a smaller initial group size should increase between-group variance. Another example is a trait affecting the propensity of the bearer to disperse from their group, which affects the frequency with which individuals from different groups mix.

We are interested here in whether such traits can evolve to increase between-group variance, and hence create the conditions for cooperation. If so, then we can provide an evolutionary explanation for how the population structures supporting cooperation arise. This is in contrast to most models of social evolution, where there is an implicit assumption that the population structure is exogenous (Peck and Feldman 1988; Peck 1992 and Avilés 2002; Avilés et al. 2002 are notable exceptions). Here, we treat such structure as endogenous and hence subject to natural selection. This then allows us to seek an explanation for how cooperation evolves without having to rely on all of the necessary structure being provided by the environment. Our treatment of population structure as an endogenous factor is analogous to the way in which mutation rates were originally assumed to be constant, but are now viewed as a factor that can co-evolve with other traits (Okasha, 2006).

In this chapter we develop a logical argument, supported by a conceptual model, for how evolution can create a selective environment favouring cooperation, by selection for a series of gradual mutations, each creating conditions more conducive for cooperation than the last. By way of illustration, we consider the individual trait of initial group size preference, that is, the number of other conspecifics with which an individual founds a group or colony. A smaller initial group size preference should lead to a greater between-group variance in behaviour if the distribution of the behavioural trait follows a binomial or hypergeometric distribution when the groups are founded. This is due to the effects of sampling meaning that each group will be less representative of the global population. The initial variance is then amplified as multiple generations of reproduction occur within the groups (Wilson, 1987; Fletcher and Zwick, 2004), leading to a greater positive assortment of behaviour in the groups that were initially of a smaller size. A smaller initial group size preference thus creates a selective environment favouring greater levels of cooperative behaviour, but how can this preference evolve? One argument suggests *a priori* that it cannot, for the following reason. Cooperative individuals should enjoy a greater relative fitness within (initially) smaller groups, since such groups will show greater positive assortment, meaning that the cooperative type experiences a greater *per capita* share of the benefits of cooperation through being clustered together. On the other hand, however, selfish individuals should enjoy a greater relative fitness in larger groups that show less positive assortment, for they will increase their share of the benefits of cooperation in such groups. Thus, if both types can evolve their initial group size preferences, then we might expect the behaviour that is in the greatest frequency in the metapopulation to modify the population structure to one that makes itself fitter. To sum up, how can a selective environment that favours cooperation be created when most individuals are initially selfish, and hence intrinsically opposed to such an environment?

Here we construct a simple conceptual model which demonstrates that this argument is fallacious, by showing that a selective gradient on the size preference allele towards a smaller size that favours increased cooperation can exist, even when the metapopulation is initially composed mostly of selfish individuals. Our model and logical argument show that the kind of selection that operates on the behavioural trait within groups is a key determinant of whether or not an adaptive gradient on initial group size (or any other trait that affects the selective environment on behaviour by affecting assortativity) exists. Specifically, we show that a gradient towards a population structure that supports increased cooperation is much more likely to exist if selection on behaviour exhibits negative frequency-dependence, leading to a protected polymorphism of behaviours. Such selection occurs in the Snowdrift (or Hawk Dove) game (Sugden, 1986; Doebeli and Hauert, 2005), and has recently been argued to occur in a number of natural environments, as discussed in Chapter 5.

We assume that the metapopulation is initially in a state where groups are founded by many unrelated individuals, hence group-level selection is a weak evolutionary force. This assumption then allows us to show how increased group-level selection can evolve through evolution of an individual trait, namely, initial group size preference. Conceptually, our starting point is that individuals initially have fitness-affecting interactions with many others, but can evolve to bring who they interact with under their control (indirectly) through the evolution of initial group size preference. This corresponds to the modelling of social interactions that are largely obligatory, such as resource consumption strategy, where an individual cannot help but affect others and in turn be affected by their behaviour. Under such interactions, some loose trait-group structure will nearly always be present (Wilson, 1975a), since interactions are rarely perfectly well-mixed. We then investigate whether this structure can evolve to one that is more conducive for cooperation. On the other hand, some other work on the evolution of social systems (such as Avilés 2002) considers social interactions that are facultative, such as communal brood care. Such models therefore start out with an assumption that individuals have no fitness-affecting interactions for the trait in question, and then show how these facultative behaviours can evolve in tandem with grouping. Both obligatory and facultative interactions are important in the evolution of social systems, and hence such models are complimentary to that presented here.

6.2 The concurrent evolution of population structure and social behaviour in a mutation model

We have a developed a simulation model where an individual's (haploid) genotype codes not only for its social behaviour, but also for a trait that affects group structure and hence the balance between individual- and group- level selection. Specifically, an individual's genotype contains two loci. The first of these loci is binary and codes for whether or not an individual cooperates within its group, that is, increases the productivity of its group at a cost relative to non-cooperative individuals within the same group [1]. The second locus codes for an initial group size preference, that is, the number of other conspecifics with which an individual founds a group. This locus influences the balance between levels of selection, because a smaller initial group size will increase between-group variance due to the effects of sampling error (assuming non-assortative group formation). Our model uses a multi-allelic integer representation for this locus, which in the simulations presented below can take values between 1 and 40. Mutations are allowed to occur separately on both of these loci: mutation on the first involves switching behaviour, mutation on the second increasing or decreasing the size preference by 1 (with equal probability).

The metapopulation in our model behaves according to Algorithm 6.1. This algorithm, along with the detailed description of each step in the following subsections, fully specifies our model.

The differences from the model in Chapter 3 are:

- Initial group size preference can now vary over an integer range, rather than being binary "large" or "small".

---

[1] As discussed in Chapter 1, this definition accords with Wilson's (1979; 1990) weak altruism and Pepper's (2000) whole-group traits, but it should be stressed that both weak and strong altruism exhibit similar dynamical behaviour when groups exist for multiple generations before dispersal. That is, both decline in frequency within each group, and hence the distinction becomes less important (Fletcher and Zwick, 2007). In particular, the arguments by Grafen (1984) and Nunney (1985) about weak altruism evolving by individual selection in randomly formed groups do not apply to multi-generational groups.

Algorithm 6.1 Aggregation and dispersal model with mutation and varying initial group size

1. **Initialise** the migrant pool with a total density of N individuals, as stated in Section 6.2.2.
2. **Form groups** from the migrant pool, according to Algorithm 6.2 below.
3. **Reproduce** within the groups for t generations, according to Equation 6.2 below.
4. **Disperse** all groups to form a new migrant pool, as stated in Section 6.2.4.
5. **Mutate** $M\%$ of the individuals in the migrant pool, as stated in Section 6.2.5.
6. Repeat from step 2 for T aggregation and dispersal cycles.

- New group size preferences can be generated by mutation.
- Groups are formed based on size preferences according to Algorithm 6.2, since not all individuals will now end up in the size of group that they specify, due to the range of group size preferences.
- The game-theoretic fitness function in Equation 6.1 can produce both directional and negative frequency-dependent within-group selection on social behaviour. Thus allows us to investigate how negative frequency-dependent selection affects the range of condiitons under which social niche construction occurs.
- Mutation can change social behaviour between cooperate and selfish, and vice versa.

This model therefore relaxes several of the simplifying assumptions made in Chapter 3, notably, that within-group selection must be directional against cooperation, and that group size preference is binary and not subject to mutation. In this chapter, we investigate what relaxing these assumptions tells us about social niche construction and when it can occur.

The next subsections describe the steps of Algorithm 6.1 in detail.

6.2.1 Migrant pool initialisation

At the start of a simulation run, we fix all individuals in the migrant pool to have the same size preference allele; this is 20 in the experiments reported here. The frequency of the cooperative allele is then set to be its equilibrium under the aggregation and dispersal process if all groups were fixed at an initial size of 20. We thus start all individuals out in the same population structure, and start the behavioural allele at its equilibrium for this structure. We then determine whether this is the *joint* equilibrium of structure and social behaviour, by allowing both initial group size and behaviour to be subject to mutation and selection over many aggregation and dispersal cycles.

6.2.2 Group formation

The metapopulation in our model is based around the aggregation and dispersal process in Chapter 3, with the addition of multiple values for the group size preference allele, and a corresponding process of group formation. Groups are formed with respect to size preference so that, to a first approximation, individuals live in a size of group that meets their preference. Subject to this constraint on the size preference allele, the assignment of the behaviour allele to groups follows a hypergeometric distribution, i.e., random sampling without replacement. Procedurally, this assignment of genotypes to groups is carried out in according to Algorithm 6.2.

Algorithm 6.2 Group formation algorithm.

```
1. Create a list of all individuals in the migrant pool.
2. Sort this list in reverse order of group size preference, such that the individuals
   with the largest size preference are at the front of the list. Within each sub-list of
   individuals with the same size preference, randomise their position in the list with
   respect to their social behaviour (cooperative or selfish).
3. Create a new group, and add the individual at the front of the list to this group.
   Remove the added individual from the list.
4. Continue adding individuals in order from the list, while the following condition is
   met: the mean size preference of the group members is less than the current group
   size. When this condition does not hold, advance to step 5.
5. If there are still individuals in the list, go back to step 3, else all groups have been
   formed.
```

Regarding step 2 in this algorithm, randomising the order of each sub-list of individuals with the same size allele means that the behaviours are assigned to groups according to a hypergeometric distribution, and not assortatively. More precisely, there is a separate hypergeometric distribution of behaviour for each value of the group size allele. Sorting the list in reverse size order handles the special case of the last group. This is because the last group will contain the handful of remaining individuals in the migrant pool. If the list was sorted in increasing order, then this last group would be small but would contain the individuals with the largest size preference. Sorting in reverse order means that the small last group contains the individuals with the smallest size preference.

Although at first appearance our procedure of forming groups may not seem particularly biologically plausible, since it contains a global sorting step, it should be remembered that what needs to be achieved is some degree of match between an individual's group size preference and the size of group that they end up in. In general, if an allele that codes for a preference for a particular kind of population structure means anything, then it must mean that individuals with that allele should (in the absence of other constraining factors) tend to end up in that structure more frequently than if they did not posses the allele. Thus, there must be some degree of assortativity on the structure allele, since individuals with a particular structural allele will want to live in the same structures. So, we should expect individuals with a certain group size allele to tend to end up in the same groups to some degree by definition, and this is what the above algorithm produces. There may be other ways of achieving this that do not rely on a global procedure, but the algorithm we use is transparent that, a) there is assortativity on the group size allele, and b) there is not assortativity directly on the behavioural allele.

6.2.3 Reproduction within groups (within-group selection)

Once the groups have been formed, we do not consider discrete individuals. Rather, we take the number of individuals with each genotype within a group to be the initial density of that genotype in the group. We then allow these densities to vary continuously as reproduction occurs within the group (as, for example, does Wilson 1975a), avoiding the need to discretise back to an integer representation of genotype density at each generation, as would occur in a strictly individual-based model. Instead, we only discretise the densities back to integers when the migrant pool is formed. Our use of densities rather than numbers within groups is analogous to the use of continuous densities in the Lotka-Volterra equations.

Asexual reproduction occurs within the groups for a number of generations, t, and causes the genotype densities to change. We use the game theory fitness function from Chapter 5 that can produce both the n-player Prisoner's Dilemma and Snowdrift games. This fitness function is repeated again below for clarity:

$$w_c = p\left(b - \frac{c}{2}\right) + (1-p)(b-c) + w_0 \qquad\qquad w_s = pb + w_0, \tag{6.1}$$

where w_c is the *per capita* fitness of all genotypes with the cooperative allele within the group, w_s the fitness of all genotypes with the selfish allele in the group, p the proportion of the cooperative allele in the group, b the benefit of cooperation, c the cost, and w_0 a baseline fitness in the absence of social interactions. Setting $0.5 < b/c < 1$ produces the Prisoner's Dilemma, $b/c > 1$ the Snowdrift game (Doebeli and Hauert, 2005). In terms of within-group selection, the size preference allele is neutral; it only changes in frequency (due to selection) when the groups compete to form the migrant pool. Thus, all genotypes with the same behavioural allele have the same fitness within any one group, and so for a given behavioural allele in a given group, the frequency distribution of size alleles that it is paired with does not change. This is shown in the difference equation below, which states how genotype densities change according to fitness:

$$n_c(g+1) = n_c(g)w_c \qquad\qquad n_s(g+1) = n_s(g)w_s, \tag{6.2}$$

where $n_c(g)$ is the density of all genotypes with the cooperative allele in the group at generation g, and $n_s(g)$ is the density of all genotypes with the selfish allele in the group at that generation. Evolution within groups according to this equation is deterministic; we do not consider genetic drift or other stochastic forces.

After every generation within groups, i.e., after Equation 6.2 has been iterated once in all groups, we rescale the size of each group to main a constant population size. This is done by normalising the size of each group in a proportionate manner, such that the sum of all group densities adds up to the original population size, keeping their size differentials after the generation of growth intact. Within any one group, the proportion of genotype densities remains the same.

In the first version of our model there is no intrinsic pressure for large or small groups, i.e., group size does not directly influence individual fitness, but only indirectly through its effect on between-group variance and hence the frequency of cooperation. Thus, the only component of selection on population structure that can potentially exist is due to social behaviour. However, there may also be other selection pressures on population structure that may oppose the component due to social behaviour. Here, as in Chapter 3, we consider the possibility of an Allee effect that favours groups of a larger initial size (Allee, 1938; Odum and Allee, 1954); for example, due to better defence against predators in larger sized groups, or access to resources that smaller groups cannot obtain (Avilés, 1999). To incorporate the Allee effect into the model, we add the following sigmoidal function to the above fitness functions:

$$\sigma_n = \frac{\beta}{1+e^{-\mu n}} - \frac{\beta}{2}, \tag{6.3}$$

where n is the *current* group size at generation g, μ the gradient (which determines how quickly the benefit tails off as the group grows), and β a parameter which determines the maximum benefit. Using this fitness function, we are able to investigate how an Allee effect favouring larger groups plays off against the advantages of greater cooperation (due to greater group-level selection) when groups are founded by fewer individuals.

6.2.4 Migrant pool formation (between-group selection)

After t generations, all groups disperse and form a migrant pool which then seeds the next generation of groups. When this migrant pool is formed, groups that have grown to a larger size will constitute a larger fraction; this is between-group selection. This between-group selection operates on the behavioural allele in the standard manner: if groups with more cooperators have grown to a larger size, then the cooperative allele may (depending on the strength of within-group selection) increase in frequency in the metapopulation (compared to the previous migrant pool), as in Simpson's Paradox (Simpson, 1951; Sober and Wilson, 1998). In addition, the frequency of the size preference allele can also change as a result of selection. For example, if groups of an initially smaller size have done better than those of a larger initial size (i.e., have grown to a greater final size after t generations, for example due to more cooperation through greater group selection), and individuals in the small groups tended to have a small allele, then the small allele will increase in frequency in the migrant pool (compared to the previous aggregation and dispersal cycle) and hence in the metapopulation as a whole.

When the migrant pool is formed, we discretise the genotype densities by rounding them to the nearest integer. The genotype densities can then be treated as individuals for the purposes of mutation and group formation.

6.2.5 Mutation

Mutation occurs after the migrant pool densities have been rescaled and discretised into an integer number of individuals. At this stage, a fraction M of individuals in the migrant pool are randomly chosen to be mutated. Of this subset of the population chosen for mutation, a fraction m have their size preference allele mutated, the remaining $1-m$ fraction their behavioural allele mutated; only one locus is mutated per individual. Mutation on our integer representation of the size preference allele occurs by decreasing its value by 1, with 50% probability, or otherwise it is increased by 1. If the size allele is already at the upper or lower limit (40 or 1 with the parmater settings used here, respectively), then it is always has 1 subtracted or added, respectively, if selected for mutation. Mutation on the binary behavioural allele (cooperative or selfish) occurs by switching to the other behaviour.

After mutation has occurred, the next generation of groups is formed as specified in Algorithm 6.2, and the aggregation and dispersal process is repeated for a sufficient number of cycles for an equilibrium to be reached.

6.3 Results

We have investigated the concurrent evolution of group size preference and social behaviour, to determine conditions under which a size preference can evolve to reinforce group-level selection. In particular, we have focussed on the effect that the type of social interaction between group members has, i.e., whether it follows a Prisoner's Dilemma or a Snowdrift game.

6.3.1 Parameter settings

In the experiments described below, we use the parameter settings in Table 6.1. We set $c = 1$, with $b = 0.9$ to yield Prisoner's Dilemma interactions, and $b = 1.1$ to produce a Snowdrift scenario (this is close to the qualitative threshold between the two games that occurs at $b/c = 1$). In all of the simulations, we follow a population of total density 1000 through 6000 group formation and dispersal cycles (preliminary experimentation revealed that this was a sufficient length of time for an equilibrium at the metapopulation level to be reached), and record the average population trajectory over 100 Monte Carlo trials. Group dispersal is set to occur after $t = 5$ generations of selection and reproduction within groups. The initial frequency of genotypes in the metapopulation is set as described in Section 6.2.1, with the population initially fixed for a size preference allele of 20. The greater rate of mutation on the size preference locus (compared to the behavioural locus) is motivated by the fact that in an explicit genetic system it may well be represented by many loci, where a change at any one of these loci could potentially affect the size preference phenotype.

6.3.2 Results with no Allee effect

We initially considered the simplified model where there is no Allee effect, meaning that the advantages of group living come entirely from interactions with cooperative individuals - there is no intrinsic advantage in numbers. Thus, the fitness of an individual is affected directly only by the composition of its group. A smaller initial group size can potentially evolve in such a situation, as illustrated by the following argument. Groups founded by fewer individuals exhibit greater between-group variance. As a result, such groups are more affected by group selection, and hence are more cooperative than those founded by a larger number of individuals. Consequently, a smaller size preference allele causes its bearers to live in, on average, groups with a greater frequency of cooperation. Therefore, individuals possessing the smaller size allele enjoy a greater frequency of cooperation, and are hence fitter than those bearing a larger preference allele. This can then create directional downward selective pressure on the size allele, whereby a smaller size preference will (largely) fix in the population, until a yet smaller preference

Table 6.1: Parameter settings used throughout Chapter 6.

Parameter	**Value**
Cost to cooperating, c	1
Benefit to cooperating (Prisoner's Dilemma), b	0.9
Benefit to cooperating (Snowdrift), b	1.1
Fraction of population mutated, M	0.01
Probability any mutation is on size preference allele, m	0.9
Generations within groups before dispersal, t	5
Value of size allele fixed in initial population	20
Smallest possible size allele	1
Largest possible size allele	40
Frequency of cooperative allele in migrant pool at initialisation (Prisoner's Dilemma)	0
Frequency of cooperative allele in migrant pool at initialisation (Snowdrift)	0.1667
Gradient of sigmoidal fitness bonus function from Allee effect, μ	0.4
Determinant of maximum benefit from Allee effect, β	1
Migrant pool density, N	1000
Number of aggregation and dispersal cycles, T	6000

arises by mutation and the process repeats. This effect implies the existence of an individual adaptive gradient towards a smaller initial group size, greater group selection, and greater cooperation.

6.3.2.1 Prisoner's Dilemma social interactions

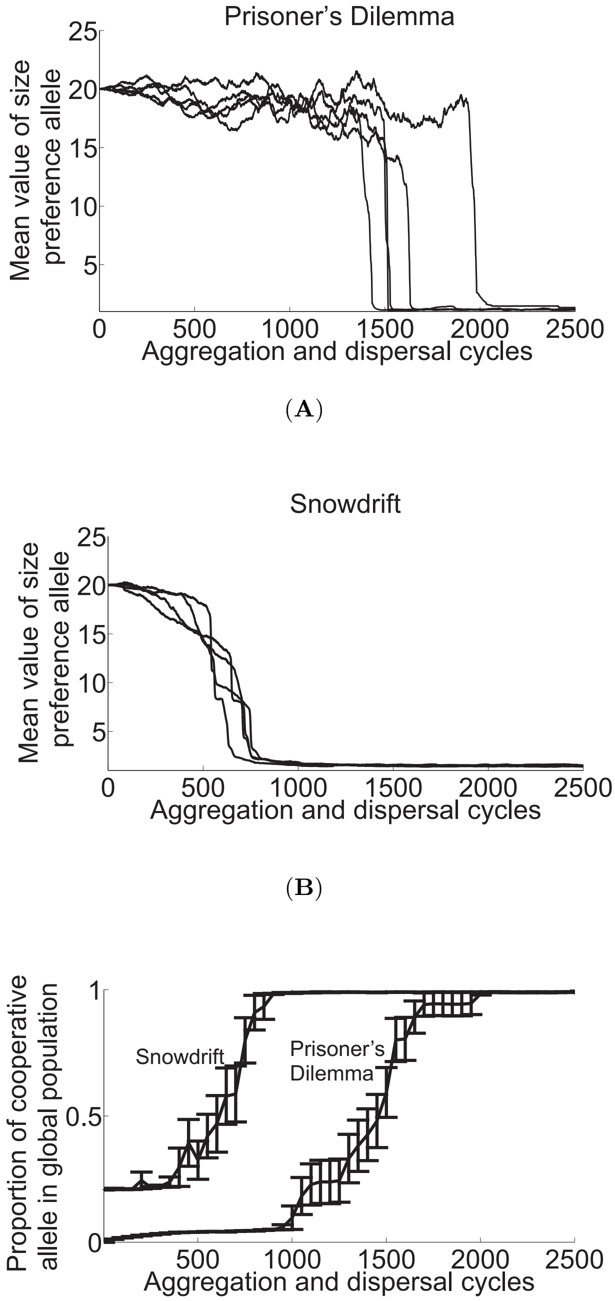
We have investigated whether such an adaptive gradient exists and whether it can be followed when smaller size alleles arise at mutation frequency. We first studied Prisoner's Dilemma interactions, and found that, from an initial value of 20, the initial size preference allele would always evolve down to the minimum possible (Figure 6.1**A**), i.e., 1, with the cooperative allele almost reaching fixation (Figure 6.1**C**). Taken at face value, this result would imply that the argument given above tells the whole story, and that the evolution of group size towards increased group selection and cooperation can proceed unhindered. However, the plot of size allele frequencies over time in Figure 6.2**A** shows that this argument is not entirely correct. This is because the allele frequencies do not gradually shift downwards as mutations accumulate, instead, size preferences spread out equally in both directions until such a time as a very small preference arises by mutation, at which point that very small size rapidly fixes in the population. This is shown by both Figure 6.2**A** and the step change in the mean value of the size allele in Figure 6.1**A**, which shows that no selection occurs until a very small size preference is reached (we show representative runs rather than the mean in Figure 6.1**A** to highlight this step change, which would be smoothed out when taking an average over multiple runs, since it happens at different times on different runs due to the stochasticity of the mutation process). The fact that the allele frequencies initially spread out in both directions shows that an adaptive gradient on this allele is not present at the start. Instead, the results imply that the size allele is neutral over most of its values, and hence changes as a result of drift rather than selection.

Why, then, is the size allele largely neutral? If a mutant smaller size preference is to be selectively advantageous, it must be the case that groups of that size are more affected by group selection, and hence their members experience a greater frequency of cooperation, than those of the current size. So, for an adaptive gradient to exist, groups of initial size $n_0 - 1$ must enjoy more cooperation than those of initial size n_0. The black line in figure 6.3 explores this under Prisoner's Dilemma interactions by calculating the frequency of cooperation in the metapopulation over a range of fixed consecutive initial group sizes, with all other parameters the same as in the previous simulation (i.e., no assortment on behaviour during group formation, 5 generations within groups between dispersal episodes, and $b/c = 0.9$). This shows that, for $n_0 > 6$, no adaptive gradient on the size allele can exist, since above this threshold moving to a slightly smaller size preference does not increase cooperation, and so size allele $n_0 - 1$ will not, on average, be fitter than size allele n_0. This threshold for the start of an adaptive gradient can also be seen in figure 6.2**A**, where allele frequencies only became concentrated around a few values once mutants with a size preference of 5 or less had arisen.

The above results suggest that an adaptive gradient on a trait increasing group selection will only exist over the very small range of parameter space where group selection is already quite effective. Thus, they suggest that we should not expect a population to evolve from a structure where group selection is weak to one where such selection is strong, for the process would seem to rely on drift from most initial conditions, and so would be unable to overcome any slight opposing selective force. However, this conclusion assumes Prisoner's Dilemma style social interactions, corresponding to directional selection against an allele for cooperation within each and every group. As discussed above, not only do such interaction dynamics represent a worst-case scenario for group selection, but there is also a growing recognition that they may not actually be the best model for many types of cooperative behaviour. Instead the Snowdrift game, which corresponds to negative frequency-dependent selection for a stable polymorphism of behaviours within each group, may be a better model. Where this is the case, we find that the results are markedly different.

6.3.2.2 Snowdrift social interactions

To investigate the effects of Snowdrift interactions, we changed b/c from 0.9 to 1.1 and then reran the experiments in the previous section. The end state was the same as in the Prisoner's Dilemma case, that is, initial group size evolved down to 1, which in turn selected for maximal cooperation between group members during subsequent group



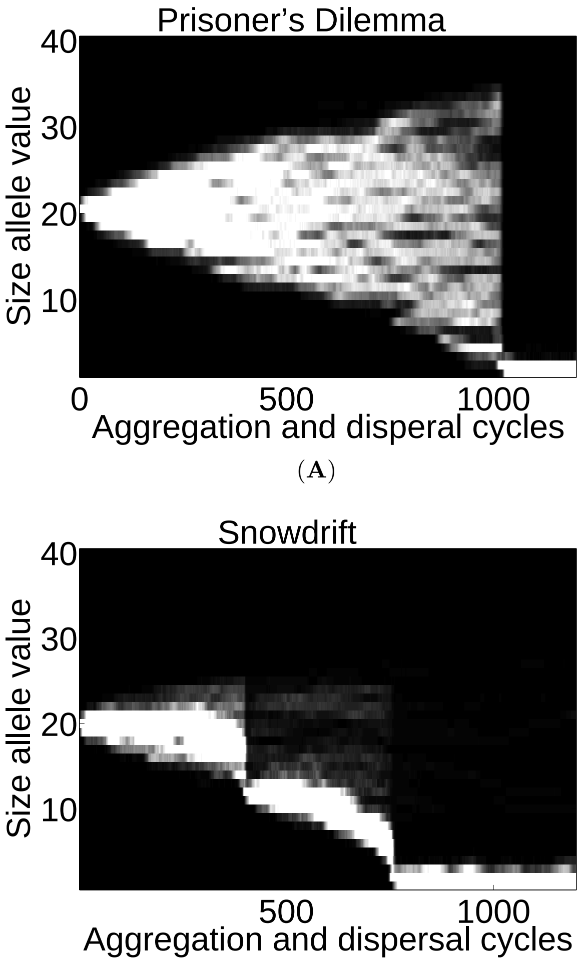
(**A**)

(**B**)

(**C**)

Figure 6.1: Concurrent evolution of initial group size preference and cooperation with no intrinsic advantage to any group size (i.e., no Allee effect). A) Mean value of size preference allele under Prisoner's Dilemma interactions in 5 representative runs. B) Mean value of size preference allele under Snowdrift interactions in 5 representative runs. C) Proportion of cooperative allele in metapopulation, averaged over 30 runs (error bars show the standard error).

growth. However, as figures 6.1**B** and 6.2**B** show, the dynamics on the transient to reach this equilibrium were very different. Rather than the size preference allele drifting in both directions, the mass of the allele frequencies showed a trend of moving downwards



(**A**)

(**B**)

Figure 6.2: (A) Size allele frequency evolution during a typical run under Prisoner's Dilemma interactions (lighter shades represent greater frequency of the allele value in the metapopulation). B) Size allele frequency evolution during a typical run under Snowdrift interactions.

from the start. In particular, the ratchet effect described in the logical argument can be seen, whereby a mutation for a smaller size preference arises and increases in frequency, until a yet smaller mutant arises by mutation - there is no evolution back towards larger preferences, even when such larger preferences were previously fixed in the population. This implied that selection was responsible for the change in size preference allele frequencies, and hence the existence of an adaptive gradient on this allele even from the initial metapopulation state.

To confirm this, we again considered whether moving from initial size preference n_0 to $n_0 - 1$ would increase the amount of cooperation its bearers experienced. The grey line in figure 6.3 shows that from the starting condition of $n_0 = 20$, decreasing initial group size by 1 would always yield some increase in cooperation due to an increased effect of group selection. The question is then: why does this occur under Snowdrift, but not Prisoner's Dilemma, interactions? We have shown in chapters 4 and 5 that negative

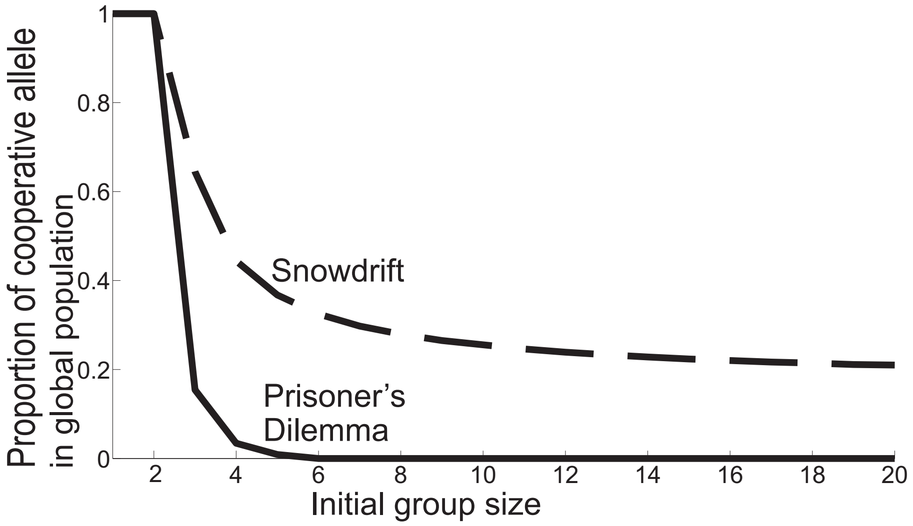


Figure 6.3: Effect of moving to a smaller initial group size on the frequency of cooperation under Prisoner's Dilemma and Snowdrift interactions.

frequency-dependent selection within groups maintains some between-group variance over a much larger range of conditions, such as initial group sizes, than if selection is directional. This is because between-group variance is proportional to the frequency of the least frequent behavioural allele in the metapopulation, and while this tends to zero under directional selection, it does not do so under negative frequency-dependence. This preservation of between-group variance means that some effect of group selection can be seen over a much larger range of conditions and, consequently, moving to a smaller initial group size increases the effect of group selection over a much larger range.

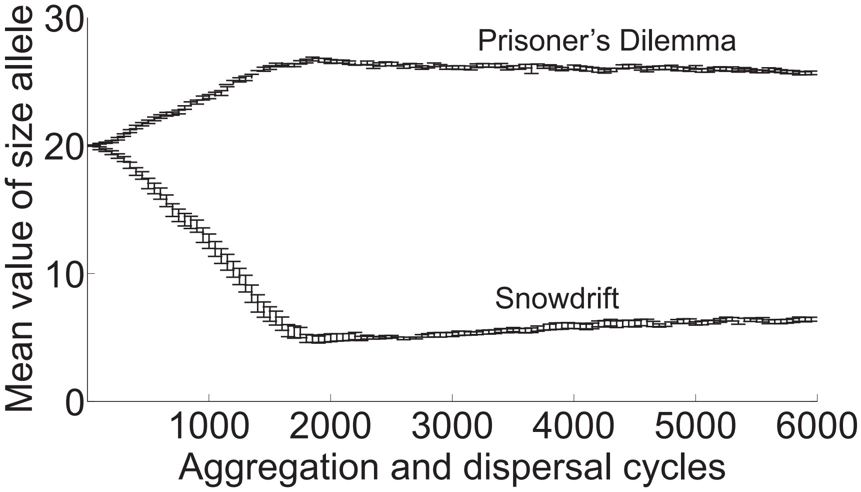
Our results therefore show that an allele coding for smaller initial group size, and hence greater between-group variance, can in fact be selected for. However, this can occur from a much wider range of initial metapopulation states if social interactions give rise to negative frequency-dependent selection, as modelled here by the Snowdrift game. These results are based on a simplifying assumption that the benefits of group living come only from the proportion of the cooperative allele, and not from any intrinsic advantage of numbers. This simplified the above experiments, by making increased cooperation the only force in the evolution of initial group size. However, in many situations a group founded by only a small number of individuals may be more vulnerable than one founded by a larger cohort. This then provides two opposing selective forces on initial group size: a short term benefit of having more members when the group is founded, versus a long term benefit of increased fitness through greater cooperation if the group is founded by fewer individuals. In the next section we examine the trade-off between these two conflicting forces by explicitly incorporating the Allee effect term of equation 6.3 into individual fitness.

6.3.3 Results with an opposing component of selection on population structure due to an Allee effect

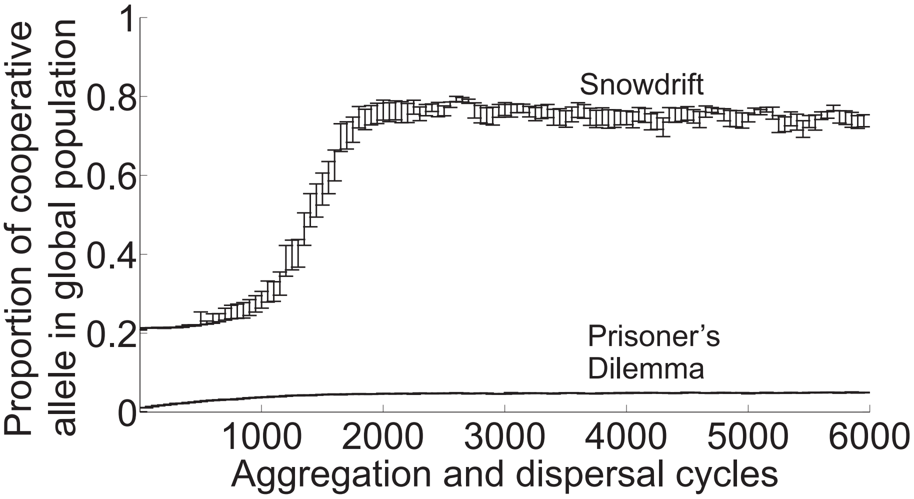
An Allee effect would favour larger numbers during the early stages of group growth, which we model using the function in equation 6.3. This benefit function is sigmoidal and hence its gradient rapidly reaches zero as group size increases - this represents the fact that, above a certain size, the advantages of number become cancelled out by the effects of increased crowding (Odum and Allee, 1954). In the absence of any group selection effects on cooperation, an initial group size of 20 or greater would be favoured using the parameter settings for the sigmoidal function described above, since this is the group size for which the gradient reaches zero. However, given that group selection effects can occur to increase cooperation, individual fitness would actually be increased at a smaller initial size that produces more between-group variance and hence stronger group selection favouring cooperation. We investigate below whether an adaptive gradient towards such an intermediate group size exists and can be followed by mutation and selection.

The results in Figure 6.4 show how the evolution of initial group size is affected by the incorporation of an Allee effect. Initial group size preference evolves downwards, and cooperation increases, under Snowdrift but not Prisoner's Dilemma types of interaction (Figure 6.4**A** and Figure 6.4**B**). This is because under Prisoner's Dilemma interactions, no adaptive gradient exists on the size preference allele (Figure 6.5**A**), in accordance with our previous results. Because of the opposing selective force towards larger groups generated by the Allee effect, the size preference allele is no longer able to drift downwards. On the other hand, under Snowdrift interactions an adaptive gradient towards a smaller size preference still exists (Figure 6.5**B**), even given the immediate advantages of numbers from the Allee effect. Thus, these results highlight the importance of an adaptive gradient, since genetic drift is unable to overcome the opposing selective force provided by the Allee effect, leaving the population at an equilibrium with suboptimal individual fitness under Prisoner's Dilemma interactions. Finally, Figure 6.4**A** suggests that the mean value of the size preference allele starts to increase again under Snowdrift interactions once the cooperative allele has reached a high frequency. This is because linkage disequilibrium evolves between the size preference and behavioural alleles, such that the selfish allele becomes associated with a large size preference, and the cooperative allele with a small preference, for reasons discussed below. It is this association of the selfish allele with a larger preference that pushes the mean value of the size allele slightly upwards towards the end of a run.

Further analysis of the results is shown in Figure 6.6. Figure 6.6**A** shows how group size preference evolves to increase between-group variance under Snowdrift interactions, thus increasing the effect of group selection. Figure 6.6**B** shows that, where initial group size evolves in this manner, linkage disequilibrium is generated between the size



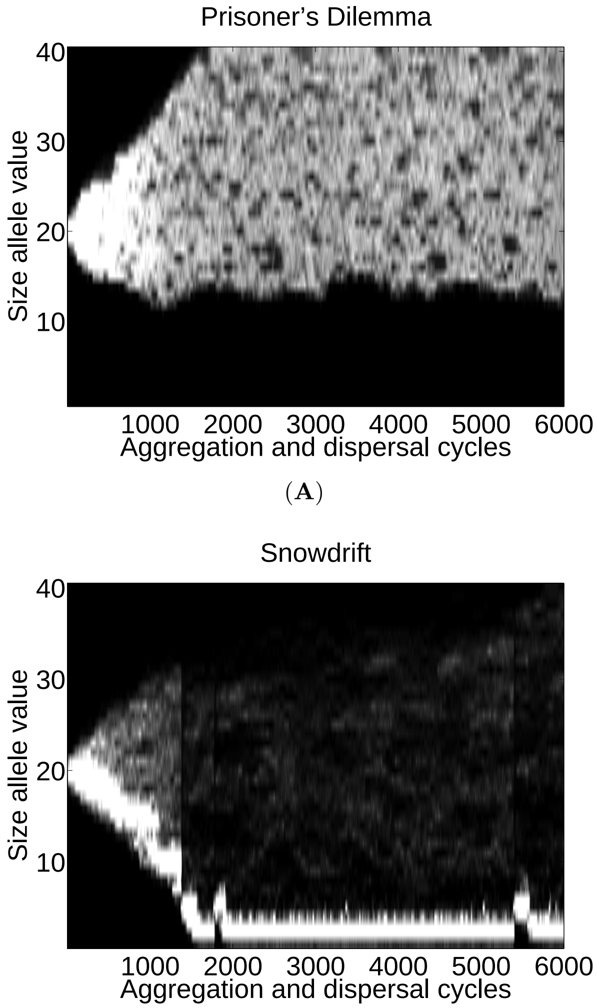
(**A**)



(**B**)

Figure 6.4: Concurrent evolution of initial group size preference and cooperation with an opposing selective force provided by an Allee effect. Error bars show standard error from 100 trials. A) Mean value of size allele. B) Cooperative allele frequency in metapopulation.

preference and behaviour loci. In particular, the cooperative allele becomes associated with a small size preference, and the selfish allele with a large preference. We found the mean linkage disequilibrium over 100 trials, calculated using Lewontin and Kojima's (1960) normalised method, to be 0.62 after 6000 aggregation and dispersal cycles (by contrast under Prisoner's Dilemma interactions the linkage disequilibrium was 0.12). Our original logical argument for the evolution of initial group size can explain why this occurs, as follows. A smaller size preference allele causes its bearers to live in groups more affected by group selection. Because of this increased group selection, increased



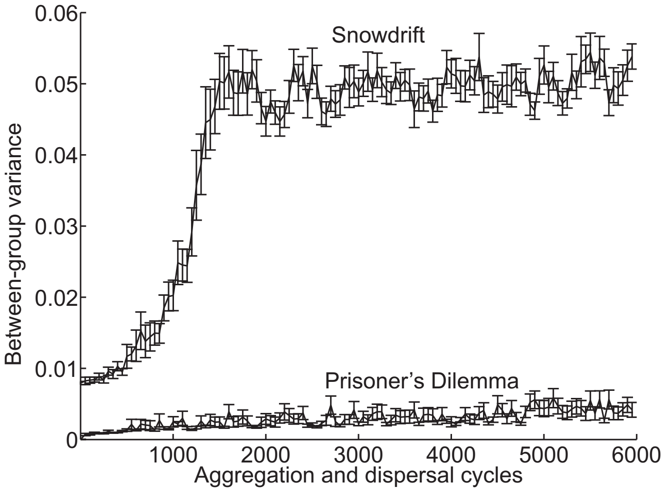
(**A**)

(**B**)

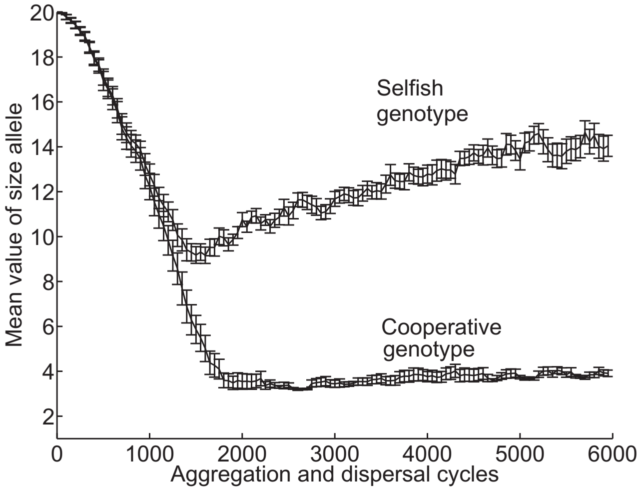
Figure 6.5: Size allele frequency evolution (lighter shades represent greater frequency of the allele value in the metapopulation). A) Representative run under Prisoner's Dilemma interactions. B) Representative run under Snowdrift interactions.

cooperation is selected for in such groups. Since bearers of a smaller size preference allele find themselves in such groups more often than bearers of a larger preference allele, an association between cooperation and a smaller preference builds up. Indeed, it is this association which allows the smaller preference allele to be selectively advantageous, since it causes its bearers to receive, on average, more of the benefits of cooperation.

It can also be seen from Figure 6.6**B** that as equilibrium genotype frequencies are reached in the metapopulation, most of the selfish alleles are associated with a size preference very much larger than the cooperative allele. This follows on from the previous argument, since selfish individuals are (largely) purged from groups that are initially very small by group selection. Hence, when the cooperative allele evolves such a preference, the selfish allele can only be maintained if its bearers have a preference for groups much less affected by group selection, i.e., groups with a large initial size. This trend then explains



(**A**)



(**B**)

Figure 6.6: Further analysis of Allee effect results. A) Evolution of between-group variance. B) Evolution of linkage disequilibrium between group size and behavioural alleles under Snowdrift interactions.

the occasional upward shift in size preference by cooperative individuals at equilibrium, seen in Figure 6.5**B**. This is because when no selfish individuals have a slightly larger size preference, the Allee effect can provide a pressure for initial size to increase. However, this is not maintained, since selfish mutants can arise and increase in frequency under such a larger size preference, causing cooperators with a smaller preference to once again outcompete cooperators with a larger preference, thereby restoring the equilibrium. In general, such cyclic dynamics rarely occurred in our model, and were only transitory, with the equilibrium genotype frequencies always restored.

We have also investigated whether the size preference allele was being used as a proxy mechanism for assortative group formation, since individuals with the same size preference tended to find themselves in the same groups. Could our results be explained entirely by this fact? To test this, we used the same group formation mechanism, except for the replacement of the size preference allele with an arbitrary numerical tag. Thus, individuals with the same tag would tend to find themselves in the same group, but the initial group size was fixed at the starting value of 20 so that the tag had no effect on group size, thereby mimicking the assortative aspects of our group formation mechanism without their effect on size. We found that in this case, no linkage disequilibrium evolved between tag and behaviour, and hence cooperation did not increase in frequency. This is because although one value of the tag would temporarily be linked more to the cooperative allele by stochastic effects, this could not be maintained under selection, since selfish individuals would also evolve the same tag, as theory predicts will happen with such greenbeard effects (see, for example, Okasha 2002 for a discussion of the vulnerability of greenbeard genes). Linkage disequilibrium could only be stable when the tag itself affected the strength of group selection, as an increase in group selection would select against selfish individuals with the same tag, as occurs here when the tag corresponds to group size preference. Thus, assortative group formation alone does not provide an explanation for our results.

6.3.4 Sensitivity to parameter settings

In this chapter we have shown how a trait affecting the strength of group selection can evolve concurrently with social behaviour, and how a sharp qualitative distinction arises between Prisoner's Dilemma and Snowdrift style social interactions. In particular, the distinction is over whether an adaptive gradient towards stronger group selection exists and can be followed by a series of mutations. The quantitative range over which this gradient occurs depends upon the range of parameter space over which an effect of group selection can be seen, and hence on the range over which some between-group variance is present. In the model presented here, the other factors apart from initial group size that determine this are the initial state of the metapopulation (size preference and cooperation allele frequencies), the degree of assortativity in group formation, the number of generations spent within groups before dispersal, and the cost-to-benefit ratio of cooperating. All of these factors matter, however, only in so far as they change the quantitative range over which an adaptive gradient towards smaller initial group size exists; the qualitative arguments presented here still hold for other settings of these parameters. In a similar manner, the parameters of the sigmoidal size benefit function representing the Allee effect determine the trade-off between the advantages of smaller initial size arising from greater group selection and cooperation, versus the raw benefits of living with more individuals. However, while the exact position of this trade-off determines the final size preference that will evolve, it does not affect the qualitative result that a smaller initial group size can evolve due to the benefits of increased group selection, nor the distinction between Prisoner's Dilemma and Snowdrift style social interactions.

6.4 Discussion

We have shown here how population structure can evolve to create the conditions for effective group-level selection and hence high levels of cooperation between group members. Crucially, this evolution of population structure can, under some conditions, occur as a result of natural selection on individual genetic traits. Thus, we have shown, both by logical argument and computer simulations, that selection on individual traits can reinforce group selection, even when group and individual selection on a behavioural trait are in opposition.

Our results complement existing work on multi-level selection, which typically either explores the range of conditions under which group selection can be effective (e.g., Levin and Kilmer 1974; Wilson and Colwell 1981; Wilson 1987; Fletcher and Zwick 2004; Peck 2004), or considers the population structure of a particular organism and determines whether group selection is likely to be important in the evolutionary history of the organism, given that population structure (e.g., Avilés 1993). However, these models do not consider that population structure can be influenced by individual genetic traits and can hence itself evolve. By contrast, we have explicitly modelled how the strength of group selection can change over evolutionary time. Furthermore, we have shown that an adaptive gradient towards the conditions for effective group selection (such as a small initial group size) can exist and be followed by a series of mutations. Thus, while previous work has shown the conditions for effective group selection, we have shown here how those conditions can themselves evolve.

It is worth pointing out that our results can also be interpreted under a kin selection framework, since the metric of between-group variance is equivalent to that of genetic relatedness (Queller, 1992b). This is particularly evident when considering initial group size, for the smaller the initial group size, the greater the relatedness of group members in subsequent generations due to descent from a smaller number of common ancestors. We have thus shown here how the evolution of genetic traits can lead to an increased relatedness between interacting individuals, and hence how a population structure that supports cooperation according to Hamilton's rule (Hamilton, 1964a) can itself evolve. Thus, while Hamilton's rule predicts the conditions for cooperation to evolve, we have shown here how a population structure that satisfies those conditions can arise by adaptive evolution. It should be stressed that restating our results in a kin selection framework in no way invalidates the group selection explanation that we have given (Queller, 1992b; Dugatkin and Reeve, 1994; Foster et al., 2006b; Wilson and Wilson, 2007). Rather, it serves to highlight the fact that kin selection is effective when there is a high relatedness between group members, and hence a high between-group variance. Thus group structure, and variance in the composition of the groups, is an essential part of either explanatory framework (Wilson, 1975a; Wade, 1980; Michod, 1982).

The arguments developed in this chapter begin with an initial population whose members interact in larger groups founded by many individuals. This assumption is valid for cases where fitness-affecting interactions between organisms are obligatory, rather than facultative. Many types of interaction are of this form, where it is a fact of life that an organism affects, and is affected by, others. Obvious examples of this type of interaction include predation (a prey organism cannot choose to simply not be affected by a predator) and consumption (when an organism consumes a resource, then less is available for other conspecifics). Indeed, two types of social interaction often considered when modelling the evolution of cooperation are a growth rate versus yield trade-off in resource consumption (Pfeiffer et al., 2001; Kreft, 2004), and public goods production (Griffin et al., 2004; Greig and Travisano, 2004; Gore et al., 2009). Both of these cases may involve obligatory interactions. For example, if organisms depend upon a shared resource then an individual cannot help but be affected by the rate and yield at which others use that resource, and vice versa. Likewise, an organism may need to utilise extra-cellular products to survive, and hence cannot help but either manufacture those products itself, or use those supplied by other individuals in the vicinity. Thus, some kind of loose trait-group structure (in the sense of Wilson 1975a, 1977), whereby individuals interact with many others, will often be present. We have shown here how this loose structure can evolve to increase the effect of group-level selection.

It is worth contrasting our approach with that of other work on the evolution of social systems, such as that by Avilés (2002). This other work typically considers social interactions that are facultative, and hence a reasonable assumption is that individuals have no such interactions at the beginning of the model. Such facultative interactions include communal brood care (Kokko et al., 2001) and coordinated group hunting (Packer and Ruttan, 1988). Because of the assumption of faculatative social interactions, this work assumes that individuals do not belong to any group at the outset, and then shows how the benefits of performing and receiving the facultative act can be a driver for the evolution of group living. By contrast, we consider obligatory social interactions, which in turn means that some type of trait-group structure will nearly always be present. We then show conditions under which this structure can evolve to increase between-group variance. Both obligatory and facultative social interactions are undoubtedly important when explaining the evolution of group living, and hence both approaches are complimentary.

In this chapter, we have modelled the evolution of initial group size to illustrate how the balance between opposing levels of selection can evolve. Roze and Michod (2001) have also considered competition between two different initial group sizes, in the form of propagule size in a model of the evolution of multi-cellularity. In their model, they showed that, if selfish mutations arise within otherwise cooperative groups, then a smaller initial propagule size can be selected for because it increases between-group variance and hence the ability of group selection to purge selfish mutants. In agreement with the results above, they showed that this could occur even when there were some intrinsic advantages of a larger propagule size, which in their case corresponded to a larger size of the adult multi-cellular organism. Our work, however, differs from theirs in several respects. The most significant of these is that they considered competition between two fixed initial group sizes which were quite far apart (5 and 1), whereas we consider how initial group size can evolve via a series of small mutations. This in turn means that we explicitly consider whether an adaptive gradient exists on initial group size and whether it can be followed by a series of mutations. On the other hand, Roze and Michod (2001) only competed two fixed sizes and hence were not concerned by the presence of an adaptive gradient. We have shown here that whether or not an adaptive gradient is likely to be present depends on the type of social interaction, and that the lack of such a gradient can frustrate the evolution of group size, even when a smaller group size would be fitter if it could be reached. Thus, our work has shown the importance of the type of social interaction, which performing static analysis by competing only two fixed group sizes does not highlight. Thus, our work is complimentary to that of Roze & Michod, for they have shown that a smaller initial group size can in principle be selectively advantageous, while we have shown by dynamical analysis conditions under which such a smaller size can and cannot be reached.

Apart from initial group size, there are other aspects of population structure that affect the balance between levels of selection and which can also be influenced by individual genetic traits. Examples include the between-group migration rate (here 0 between dispersal episodes, and 1 at the dispersal stage), the number of generations within groups prior to dispersal (here 5), and the degree of assortativity at the behavioural locus during group formation (here 0). Traits affecting these factors could also evolve alongside social behaviour in a manner analogous to that modelled in this paper. Of particular interest is the dynamics that could occur when more than one of these population structure modifying traits evolves concurrently; we intend to make this the subject of a future study.

It is also worth discussing how the assumptions of our model influence the results presented here. The metapopulation structure that we have modelled is based on the classic Haystack model, and represents organisms that live on ephemeral resource patches which last for a number of generations. While this fits the biology of some organisms particularly well, it is of course the case that other population structures, such as “stepping stone” or “island” style migration between non-ephemeral patches (Wilson, 1992), may be a closer model for other organisms. However, we have used the Haystack model here to capture the abstract logic of multi-level selection arguments, in a manner similar to the way in which it was used by Sober and Wilson (1998) to explicate the general tension between opposing levels of selection. Logical argument suggests that the qualitative conclusions of our study should be little affected by the exact nature of the organism's population structure. In particular, negative frequency-dependent selection, as occurs in the Snowdrift game, should always allow for the possibility of some between-group variance to be maintained, since no behaviour can be driven extinct. Therefore, such interactions should still be more conducive to providing an adaptive gradient towards increased group selection. Testing this hypothesis through the application of our model to other metapopulation structures provides an interesting avenue of future research.

Another assumption of our model is that of asexual reproduction. This is important, in so much as linkage disequilibrium is generated between the size preference and social behaviour alleles in our model. An interesting future study would be to determine how sexual reproduction, which has the potential to disrupt linkage disequilibrium, would affect our results. However, since the linkage disequilibrium in our model is generated by selection, in particular by a smaller size preference allele causing increased cooperation amongst its bearers to be selected for, we hypothesise that sexual reproduction should make only a quantitative difference to our results. Finally, a global carrying capacity is imposed in our model by capping the total population size, rather than by limiting the number of groups (which would correspond to a limit in the number of resource patches upon which groups live). Limiting the number of groups may create an upward pressure on initial group size, resulting in a trade-off with the effect of group selection, in a manner similar to the trade-off arising due to the Allee effect. Either or both of a carrying capacity limiting the number of individuals, or limiting the number of groups, may occur in natural populations.

When considering the wider implications of our results, they reinforce the need to select an appropriate model for the particular type of cooperative behaviour occurring in any given ecological scenario. In the past the Prisoner's Dilemma has typically been assumed to be a suitable model for many types of cooperative act, particularly in multi-level selection models where it is usually implicit in the within-group fitness equations (Fletcher and Zwick, 2007). More recently, however, the validity of Prisoner's Dilemma interactions has been questioned for a number of ecological behaviours, including a variety of public goods scenarios. For example, it can be the case that cooperators receive a greater *per capita* share of the public good that they produce. Significantly, this can occur even in the absence of positive assortment, since a cooperator may, for example, be able to internalise a fraction of the good that it produces, such that only it benefits, as has been shown to be the case with digestive enzyme production in yeast (Gore et al., 2009), and antibiotic resistance in bacteria (Dugatkin et al., 2003). If the *per capita* benefits of the public good decrease with increasing frequency, then this will lead to negative frequency-dependent selection on production behaviour, exactly as we have modelled here with the Snowdrift game. This dynamic can occur even when interactions are freely-mixed, and hence can operate within single groups[2] . Crucially, this still creates a situation where individual- and group-level selection are in opposition, since the group will benefit from a higher level of the public good than can be maintained under individual selection (see also Dugatkin et al. 2005 for a study of this phenomenon in single generational trait-groups). In chapters 4 and 5 we have shown that when social behaviour gives rise to this kind of dynamic, then between-group variance on the behavioural trait can be maintained over a much larger range of parameters. This is because a polymorphism of behaviours allows for some between-group variance to be regenerated when new groups are formed. Here we have shown that this maintenance of between-group variance, and hence of group selection, is crucial for the presence of an adaptive gradient on a group-structure modifying trait. Thus, such behavioural dynamics allow for the adaptive evolution of population structure in support of group selection.

6.5 Conclusion

Most models of social evolution assume that population structure is fixed, or changes only as a result of external disturbances. However, organisms across all taxa can in fact modify their population structure to varying degrees, for example, by influencing assortativity on behaviour, migration rate, or group size. We have shown here that explicitly incorporating such effects into a model of social evolution can have important consequences. In particular, we have shown that individual modification of population structure can create the conditions for effective group selection and, consequently, cooperation. Thus, while many previous models have shown the conditions under which cooperation can evolve, we have demonstrated here how those conditions can themselves arise as a result of individual adaptation.

---

[2] Negative frequency-dependent selection can also operate within multi-species groups, as modelled by coexistence of competing species in Lotka-Volterra competition equations, for example (see Powers et al., 2008).

Chapter 7

Social Niche Construction and the Major Evolutionary Transitions

A change in population structure is central to most of the major transitions in evolution. That is, the population structure of the pre-transitional individuals changed such that they went from a solitary existence to living in highly cooperative groups. Some part of this structure can sometimes be provided by the physical environment. Consider, for example, the hypothesis by Michod (1983) (and advanced by Maynard Smith and Szathmáry 1995) that the first replicators were randomly washed onto a "rocky shore", and that being on this shore created spatially localised interactions as compared to existing in a well-mixed solution. The replicators were then periodically washed away and freely-mixed, before being returned to the surface. In effect, a trait-group structure was created "for free" by the environmental conditions (Michod, 1983), and this allowed cooperation to evolve. Our contention, however, is that such a degree of population structure need not always be provided by the physical environment. Rather, the social niche construction thesis developed in this dissertation predicts that it can be created by the evolving individuals themselves. Moreover, we have argued that the component of selection on population structure due to social behaviour must be in the direction of creating conditions for greater cooperation. Thus, social niche construction will act as a driver towards creating the kinds of cooperative groups seen in the major transitions. Of course, how far it succeeds in this depends upon the degree of opposing selective forces provided by other ecological factors, such as crowding and resource competition. Generally, however, we argue that social niche construction provides a new perspective for understanding the process of a transition, for it can provide an adaptive explanation for how individuals create the highly cooperative groups that develop in many of the transitions.

We have provided a general logical argument for how social niche construction will create a component of selection on population structure in favour of cooperation in Chapter 1. We have also illustrated this argument with models of the evolution of the example trait of initial group size preference. Here, we discuss examples of other kinds of social niche construction, and their applicability to the major transitions.

7.1 Social niche construction and the fraternal transitions

The models in this thesis have been mostly concerned with the trait of initial group size preference, that is, the number of individuals that found a group or colony. A key feature of many of the fraternal transitions is, as discussed in Chapter 5, founding of the group by one or a few individuals. For example, most multicellular organisms develop from a single fertilised egg (Wolpert and Szathmáry, 2002). Why is this the case? Surely, if multicellularity has an adaptive advantage, there must be some benefits in number, i.e., in being in a group size greater than 1. In particular, the Allee effect (Allee, 1938) suggests that very small groups should be more vulnerable to extinction, for example, due to increased predation or difficulty in accessing resources (Grosberg and Strathmann, 1998). So, it is not obvious why cell-groups are indeed founded by only 1 individual (Grosberg and Strathmann, 1998; Wolpert and Szathmáry, 2002). Moreover, such a bottleneck in development is common to several of the transitions (Maynard Smith and Szathmáry, 1995; Queller, 2000), for example, eusocial insect colonies develop from one or a few queens (Maynard Smith and Szathmáry, 1995; Queller, 2000). The fact that this is a general pattern adds further support to the claim that it is an adaptation. Indeed, a developmental bottleneck is seen by Godfrey-Smith (2009) as a hallmark of an evolutionary individual

Michod and Roze (2000); Roze and Michod (2001) have shown that a smaller initial cell-group size can be selectively advantageous because selfish mutants that arise during development can be more effectively purged. The spread of such selfish mutants is favoured by within-group (within multicellular-organism) selection (Pál and Papp, 2000; Roze and Michod, 2001). Roze and Michod (2001) argue that the greater between-group variance that results from development from a single cell increases the effectiveness of group selection in purging these mutants. This is because the presence of selfish mutants lowers cell-group functionality and hence the success of the group in competition with others. In general, the deleterious mutational load experienced by cell-groups increases with the number of cells that found the group (Kondrashov, 1994).

It is insightful to consider the two different perspectives from which authors have considered single-celled development. Authors such as Roze and Michod (2001) emphasise the role of single-celled development in increasing the efficacy of between-group selection. Indeed, Michod (1999) sees the suppression of within-group selection, and the transfer of all variance in fitness to the group level, as the hallmark of a major transition. On the other hand, authors such as Maynard Smith and Szathmáry (1995), Pál and Szathmáry (2000) and Queller (2000), emphasise the perspective of genealogical relatedness rather than between-group selection. For them, the fact that multicellular organisms develop from a single cell means that there is no evolutionary conflict of interest between the cells that make up the group, for they are genetically identical. Because of this, they argue that selfish mutants would not be selectively favoured. This provides a nice example of the difference between kin and group selection perspectives of looking at the same population structure. Kin selection emphasises that there is no conflict between cell-group members because they are genetic clones, whereas the group selection perspective emphasises that this shared interest is realised through competition with other cell-groups. In particular, we would argue that the group selection perspective is particularly valuable in explaining the *maintenance* of the integrity of the multicellular organism in the face of selfish mutants, since deleterious mutations are purged by cell-groups with a lower mutational load outcompeting those with a higher load (Roze and Michod, 2001). This seems particularly relevant before a germ-soma separation evolved, since selfish mutants could then be passed down through the lineage to offspring, whereas after a germ-soma separation they can only reduce the survival and / or fecundity of the one cell-group in which they reside. On the other hand, the kin selection perspective seems particularly useful in understanding the *origin* of multicellularity (Maynard Smith and Szathmáry, 1995; Pál and Szathmáry, 2000; Wolpert and Szathmáry, 2002), since the fact that the first multicellular organisms may have developed clonally (i.e., with a relatedness of 1) meant that selfish behaviour was much less of an issue. A reproductive division of labour could evolve, also due to clonality, because one cell could carry the genetic information of the entire group, and so not all cells would need to reproduce in order to pass on their genes, hence some could specialise in vegetative functions (Queller, 1997; Michod, 1999). By contrast, in multicellular forms that develop by aggregation of unrelated cells, such as cellular slime moulds (Pál and Papp, 2000), selfish behaviour is much more of an issue and frustrates the functional integration of the cell-group to a much greater extent.

The *origin* of a unicellular bottleneck in the development of multicellular organisms needs explaining. In particular, Michod and Roze (2000) suggest that the ancestors of metazoans were colonies of choanoflagellates that likely reproduced by fragmentation, that is, by breaking off a propagule of several cells. Hence, the initial size of the cell groups is postulated, under this theory, to be greater than 1. If this natural history hypothesis is correct, then the evolutionary origin of a cell-group being founded by a single individual needs explaining, for that population structure was not necessarily provided "for free" by the environment. Social niche construction can provide an adaptive explanation for the origin of a unicellular bottleneck in multicellular development, by showing how a series of mutations that decreased initial group size could be selectively advantageous. Our model of the evolution of initial group size in Chapter 6 is closely analogous to this. For example, where there is no Allee effect, and hence no opposing selective force on group size, we found that individuals evolve to live in groups founded by a single individual. We showed (in the case of the Snowdrift game) the presence of a selective gradient towards this, illustrating that it can be adaptive. Because there were no other components of selection on group size, it was adaptive precisely because it increased between-group variance and hence cooperation. The flip side of this is that it reduced within-group variance, and hence reduced selection favouring selfishness, ultimately eliminating it entirely, exactly as occurs in multicellular development from a single cell. More generally, our logical argument in Chapter 1 shows that in the absence of other opposing selective forces this must occur, since living in an environment with greater cooperation raises fitness, by definition, and so individuals who live in those environments must be favoured over those that live in environments selecting for selfishness, all other factors being equal. We also investigated the presence of an Allee effect, and found that under the model parameters studied, this pushed the equilibrium initial group size upwards from 1. This has not occurred in metazoans, for two possible reasons. Firstly, the ecological parameters may be such that the Allee effect does not outweigh other advantages of single celled development, be they purging selfish mutants (Roze and Michod, 2001) or allowing for a developmental programme (Wolpert and Szathmáry, 2002). Secondly, metazoans may have evolved mechanisms for overcoming the ecological disadvantages of single celled reproduction, for example, by increased parental care of offspring (Grosberg and Strathmann, 1998).

One important difference between the model in Chapter 6 and the way that metazoan development occurs is that we have modelled group formation by the aggregation of unrelated individuals. By contrast, the early ancestors of metazoans founded groups by propagules of *related* cells, all descended from the same parent group. Thus, within-group variance, and hence selection for selfish behaviour, would initially be much greater in our model. In that sense, the initial conditions in our model are less favourable for cooperative behaviour than they likely were at the start of the transition to multicellularity. Nevertheless, we would argue that our model highlights the general argument that cooperation can drive the evolution of population structure in scenarios like the transition to multicellularity, by creating an adaptive gradient on structure-modifying traits. In this transition, social niche construction takes the form of selection against selfish mutants driving the evolution of a smaller propagule size, ultimately leading to unicellular development. Despite the fact that all group members were related by descent in the ancestors of metazoans, selfish mutations still occurred as the cell-groups grew, and the adaptive advantages of purging them could create selective pressures for social niche construction.

7.2 Social niche construction and the egalitarian transitions

The start of a fraternal transition appears to be related individuals that lived close together forming groups. In some sense, then, social niche construction involves tinkering with this population structure that was already provided by spatial collocation. In the egalitarian transitions, however, the role of social niche construction is perhaps more pronounced. This is because the pre-transitional individuals were of different types, and hence not related (Queller, 1997). Thus, a population structure already giving high between-group variance, and low within-group variance, was not provided by the environment, as it is in the fraternal transitions. This is why, for example, Queller (1997, 2000) views control of conflicts as the greatest hurdle in the egalitarian transitions. From this point of view, there is more for social niche construction to do in such transitions.

The trait of initial group size that we considered above involved the indirect evolution of associations between individuals. That is, by evolving a smaller initial group size, individuals increased the genetic homogeneity of their group. In so doing, they began to bring an aspect of their selective environment under their own control, namely the genetic context in which social interactions took place. A more extreme form of this would be the evolution of specific associations (Watson et al., 2009; Mills and Watson, 2009), whereby in an egalitarian transition individuals could evolve associations with particular other species, ensuring that their lineages carried on interacting down through the generations. For example, when genes evolved to form chromosomes, they ensured that the other genes that they were co-adapted with would be passed down with them to the next generation (Maynard Smith and Szathmáry, 1995). In this way, they stabilised their selective environment, allowing for greater co-adaptation. In general, stabilising the selective environment by evolving some kind of associations should increase the fit between organisms and their environment, for two reasons. First, if two types are co-adapted, then increasing their frequency of interaction increases how useful a particular co-adaptation is, in the sense that it can be used more frequently. Second, once the selective environment is stabilised, greater co-adaptation can occur, even to the point that organisms can evolve adaptations that mean they depend on the presence of each other.

7.3 What constrains social niche construction?

The type of social niche construction that can occur is constrained by the type of evolutionary game that group members play (and more than one game may be appropriate, see Chapter 5 for a discussion of how both coordination of function and a "Tragedy of the Commons" may need to be modelled by appropriate games in the egalitarian transitions). In particular, the evolution of direct associations between individuals is problematic in single ESS games, since all types of individual are fitter if they interact with a cooperator (this is true in both the Prisoner's Dilemma and Snowdrift games). Thus, in general we would not expect directly assortative interactions to be stable in such games. For example, we would not expect cooperators to be able to evolve a marker trait to recognise and form groups with each other, since there would be selection pressure for selfish individuals to evolve that same marker. Moreover, selfish individuals with that marker would then be fitter than cooperators with the marker. This is exactly what we saw in Chapter 6 when group size preference was replaced with an arbitrary marker, which was unable to create a population structure that sustained cooperation. Social niche construction therefore takes more indirect routes to stabilising the selective environment in such cases. In particular, the evolution of increased assortativity occurs by indirect means, such as by evolving group size or migration rate. We suggest that these are the types of population structure-modifying traits that are likely to be important in the fraternal transitions.

On the other hand, in games with multiple ESSs we would expect the evolution of direct assortativity to be stable (see, for example, Snowdon et al. 2009). For example, in a coordination game each type is fitter when it interacts with other individuals of the same type, so there is no selection pressure for trying to break the other type's marker system. Similarly, in the Stag Hunt game stag hunters do well when paired with each other, but a hare hunter is no fitter paired with a stag hunter than paired with another hare hunter, so again there is no pressure for selfish individuals to pair with the other type. In general, if there are multiple ESSs then it will be in the selective interests of a genotype to increase its frequency of co-occurrence with other genotypes that are also present in that same ESS (Watson et al., 2009). We take assortativity in a mixed ESS game to mean types preferentially interacting with the other types with which they co-occur at an ESS. This co-occurrence can be between individuals of different species, as in symbiosis and the egalitarian transitions.

It is important to remember that we posit social niche construction as an adaptive process. Therefore, at each stage it must be in the immediate selective interests of the individual to modify their population structure in a particular way. For example, we do not suggest that one type would evolve an association with another unless it was immediately advantageous. Thus, there must exist an adaptive gradient on a population structure-modifying trait that can be followed by mutation and subsequent selection. For such a gradient to exist due to social behaviour, as posited in this dissertation, three conditions must be met:

1. Individuals that create the structure must tend to live in it more frequently than individuals that do not create it.

2. The modified structure must cause the individuals that live in it to enjoy greater cooperation.

3. The increase in fitness from this greater cooperation must outweigh any fitness loss that occurs in the new structure, for example due to ecological factors such as the Allee effect.

When these three conditions are met, population structure will adaptively be modified to increase cooperation, assuming that a heritable structure-modifying trait exists. We suspect that the third of these conditions will often be the stumbling block in nature, since other ecological pressures apart from those on social behaviour may be too strong to be the overcome by the benefits of increased cooperation. Whether or not this is the case requires detailed ecological knowledge and needs to be verified on a case by case basis.

7.4 Conclusion

In many of the major transitions, the population structure of individuals changed so that such extreme forms of cooperation as a reproductive division of labour were selectively favoured. The major transitions are therefore not just about how individuals adapt to their current social environment, but also how that environment itself changes. Thus, if we want to provide an adaptive explanation for the transitions then we need to also provide an adaptive explanation for how the population structure changed to support such a high degree of cooperation. Therefore, conventional models of social evolution that just explain how organisms adapt their social behaviour to their current population structure cannot tell the whole story. It is, therefore, essential to take account of the evolution of population structure as well as social behaviour.

In this dissertation, we have argued that social behaviour can drive the evolution of population structure. In the case of the fraternal transitions, as modelled by single ESS games, we have argued that the modification of population structure will not take the form of direct assortative interactions, for that would be vulnerable to cheating on any assortative marking trait. Rather, assortativity can increase by indirect means, such as a reduction in initial group size or migration rate. We have suggested, in agreement with Michod and Roze (2000), that such a reduction in initial group size occurred during the transition to multicellularity, and was selectively advantageous because it allowed selfish mutants to be more effectively purged through stronger cell-group selection. Our model in Chapter 6 provides a general illustration of how an adaptive gradient can exist on initial group size that can be followed by a series of small mutations. In the case of the egalitarian transitions, where we have argued a multi ESS game is a better model of the social interactions, then direct assortativity can be stable. Hence direct associations, such as the evolution of linkage between single genes, can evolve.

Chapter 8

Discussion and Conclusions

The question of how cooperative behaviours can evolve has troubled evolutionary biologists since Darwin (1871). We know that population structuring can make cooperation selectively advantageous for the individual, by clustering cooperators together such that they enjoy the benefits of each others' actions. Such structuring is fundamental to the frameworks of kin (Hamilton, 1964a; Wade, 1980; Michod, 1982), and group (Wilson, 1975a, 1980; Hamilton, 1975), selection (Eshel and Cavalli-Sforza, 1982; Michod and Sanderson, 1985; Sober, 1992; Fletcher and Doebeli, 2009).

What has been left largely unaddressed by previous works, however, is how these population structures are created. In particular, the typical model of the evolution of cooperation shows how individuals adapt their social behaviour to fit the structure that they find themselves in. For example, if there is a certain degree of assortment of cooperators then the model will show that cooperation is individually adaptive, otherwise it will show that selfishness is adaptive. Such models make an implicit assumption that the population structure does not change during the course of evolution, and hence they cannot explain the origin of a structure that supports cooperation. As a result, we have argued that such models cannot really explain the evolutionary *origin* of cooperation, for surely we need to understand how the selection pressures towards cooperation were themselves created. If population structuring creates the selective pressures for cooperation, then this must involve a change in population structure. The need to explain how a selective environment that favours cooperation is created is particularly evident when considering an adaptive explanation for the major transitions in evolution, since a change in selection pressures that favoured greater cooperation must have occurred.

Most models of social evolution, then, consider the equilibrium frequency of cooperation in a given population structure. By contrast, we have explicitly considered the joint equilibrium of social behaviour and population structure. We can thus explain why organisms live in an environment that supports a particular level of cooperation, rather than in an environment that supports more or less cooperative behaviours. This is a question that models which treat population structure as exogenous cannot help to answer, unless they posit that population structure is entirely the product of physical environmental features. Some structure can sometimes be provided "for free" by the physical environment. For example, it has been suggested (Michod, 1983) that the first replicators found themselves regularly washed onto a rocky shore, which created a trait-group structure (*sensu* Wilson 1975a, 1980) that allowed cooperation to evolve. However, it is also widely appreciated that individual genetic traits can modify population structure (e.g., Orians, 1969; Emlen and Oring, 1977; Rodman, 1981; Koenig, 1981; Johnson and Gaines, 1990). Hence, the structure can be subject to selection and change in an adaptive manner, for individuals can create a structure that increases their own fitness. We have argued here that social behaviour can exert an indirect selection pressure on structure-modifying traits, causing individuals to adaptively modify their population structure to support greater cooperation (Powers et al., 2007; Powers and Watson, 2009). Thus, we are able to provide an adaptive explanation for how the structures supporting cooperation arise, rather than having to rely on them being provided by chance by the physical environment.

Moreover, we have argued that any selection pressure on population structure resulting from social behaviour must be in the direction of creating a structure that increases cooperation; that component of selection could not favour the creation of a structure that increased selfish behaviour. This was shown by the logical argument presented in Box 1.2. Our argument holds that linkage disequilibrium will evolve between behavioural and structure-modifying traits, if the following premises and assumption are satisfied (these are the premises and assumption to our logical argument in Box 1.2):

Premises

1. Population structure affects selection pressure on social behaviour, as proved by Hamilton (1964a, 1975) and Price (1972).
2. Structure A selects for greater cooperation, structure B for greater selfish behaviour (made logically possible by premise 1).
3. Individuals have a heritable trait that produces population structure A or B. Individuals also posses a heritable social behaviour (cooperative or selfish).

Assumption Individuals with trait A tend to find themselves in structure A more frequently than individuals that do not posses the trait. Likewise with trait and structure B.

We argued that our assumption will be valid whenever the structure-modifying trait creates a structure differential between bearers and non-bearers. What then matters is that the trait causes its bearers to experience *relatively* more cooperation. If the trait affects the structure of bearers, then this will occur when the new structure selects for increased cooperation. Conversely, if the trait affects the structure of non-bearers, then this will occur if the new structure of the non-bearers induces them to be more selfish. On the other hand, if the trait affected the structure of all population members, then it would be selectively neutral. In general, we would not expect a trait to modify the structure of all population members. For example, the trait could be the amount of extracellular matrix produced by a bacterium in a biofilm (Flemming et al., 2007). This would affect that bacterium, and its neighbours, but because a biofilm is not freely-mixed, it would not affect all individuals. Another example concerns aspects of the mating system in sexual populations, such as a preference for inbreeding (Wade and Breden, 1981; Breden and Wade, 1991), or monogamous relationships (Peck and Feldman, 1988). Individuals with an allele coding for such a preference would practise inbreeding or monogamy, whereas those without the allele would not. Thus, such traits would differentially affect the population structure of self. In the absence of interactions being freely-mixed, we would suggest that most structure-modifying traits are likely to have localised effects, and to hence create the structure differential required for our logical argument.

Chapters 5 to 6 investigated the second premise in detail. This premise requires there to be a greater between-group variance in social behaviour under structure A than under structure B. If new population structures arise by small mutations on structure-modifying traits, then a small change in say, initial group size, needs to increase the between-group variance. We found that where within-group selection leads to a competitive exclusion of social behaviours, then any possibility for between-group variance tends to eliminated over a large range of conditions, since one behaviour is driven to fixation. On the other hand, if there is a protected polymorphism of behaviours within groups, then the potential remains for some between-group variance to be generated when the groups are mixed and reformed (Powers et al., 2008). In particular, between-group variance is proportional to the least frequent behaviour in the global population (Wilson, 1980, 1983a), and under competitive exclusion dynamics, as given by directional within-group selection, this can be zero for much of the parameter space. On the other hand, a protected polymorphism prevents this determinant of the between-group variance from going to zero. Such a polymorphism is provided by negative frequency-dependent selection, as can be modelled by coexistence in the Lotka-Volterra competition equations (Chapter 4), and the Snowdrift or Hawk Dove game (chapters 5 and 6). We found that where such polymorphisms occur, some effect of group selection in increasing cooperation could be seen over a much larger range of parameter space. Consequently, a small change in structure can increase cooperation over a much larger range, and hence premise 2 can be satisfied from a much larger range of initial conditions (Powers and Watson, 2009).

8.1 The role of groups in evolution

The inclusive fitness, or kin selection, explanations for the evolution of social behaviour that are favoured by many evolutionary biologists do in fact make essential appeal to group structure (Wilson, 1975a; Wade, 1980; Michod, 1982; Queller, 1992b). This is because they require that the benefits of cooperation fall preferentially on other cooperators, and this cannot occur in a freely-mixed population. Thus, they make appeal to interaction-, or trait-, groups – subsets of the population in which social interactions take place. However, this group structure is not made explicit in the models but is instead implicit in the calculation of the relatedness term. In particular, relatedness is defined as the deviation in the frequency of cooperation in the actor's group from the global mean frequency of cooperation[1] (Grafen, 1985; Frank, 1998). It is thus equivalent to the concept of between-group variance (Queller, 1992b). Despite this, many authors have sought to sideline the role of interaction groups in social evolution, regarding them as merely the background context in which individual selection occurs, rather than as being targets of selection in their own right (Maynard Smith 1998; Gildenhuys 2003; West et al. 2007a advocate this position, and Dugatkin and Reeve 1994, Sterelny 1996 and Kerr and Godfrey-Smith 2002 provide a detailed philosophical justification of this view without advocating that it always be taken). The tendency is to view purported models of group selection as merely showing frequency-dependent selection on individual behaviour (Sterelny, 1996; Maynard Smith, 1998). The argument goes that cooperation is individually advantageous when the group structure provides assortment of behaviours. That is, cooperation has a frequency-dependent advantage if individuals with the same behaviour tend to find themselves in groups together. Thus cooperation evolves, on this account, by individuals adapting their social behaviour to the group structure that they find themselves in. Consequently, on this understanding group structure plays the same role in natural selection as any other characteristic of the environment: it is just something that selection causes organisms to adapt to. This is particularly vivid in an example that West et al. (2007a) give of the public goods game being played in groups that compete with each other. They imply that group selection is not involved because it is individually advantageous for an organism to share resources with group members, if doing so increases the productivity of the group and hence in turn of the individual. Similar arguments of cooperation as being an individual adaptation to group structure, but with no recourse to group selection, are frequent in the literature (e.g., Nunney 1985, 1998; Griffin et al. 2004; Wild et al. 2009; Bryden and Jansen 2010).

One line of argument in the literature that has tried to revive the role of groups has been to emphasise the distinction between type 1 and type 2 group selection, as discussed in Chapter 2. Okasha (2001, 2006) and others have stressed that the concept of type 1 group selection is that group character causally influences individual fitness. Thus, its

---

[1] In general, it is the deviation in allele frequencies at a focal locus in the actor's group compared to the allele frequencies at that locus in the global population.

explanatory target is an individual-level trait, such as cooperation, and so its results can necessarily be accounted for in terms of individual fitness (by contrast, type 2 selection is the process by which a group-level trait increases in frequency in the metapopulation of groups Damuth and Heisler, 1988; Okasha, 2006). However, it is argued that this is not to deny the causal role that group structure plays in type 1 models, in particular, that groups contribute different numbers of individuals into the population, and that this is necessary for social behaviour to evolve (assuming the trait-group definition of a group advocated by Wilson 1980; Sober and Wilson 1998 is used). Furthermore, it is suggested that since nearly all purported models of group selection have been of type 1, even ones that individualists such as Maynard Smith would agree are group selection (e.g., Maynard Smith 1964; Levin and Kilmer 1974), it would be strange not to view this process as group selection (Okasha, 2001, 2006). However, this has not settled the debate between group and individual selectionists, since a description of groups as providing the context for frequency-dependent individual selection is still available and is apparently just as satisfactory (Dugatkin and Reeve, 1994; Sterelny, 1996; Kerr and Godfrey-Smith, 2002). Thus, the argument that type 1 group selection really *is* a kind of group selection, and that authors of models making appeal to group structure must recognise that they are making appeal to group *selection*, does not wash with individualists. On the other hand, such individualists would view type 2 models as being instances of “real” group selection. However, we have argued that type 2 models can seem unsatisfying if we just assert that groups should be treated like individuals. That is, such models beg the question of how the groups came to be cohesive enough for this to be the case. Thus, to an individualist, type 1 group selection is explained away by individual selection, and type 2 group selection can seem mysterious if one cannot explain how the groups came to be cohesive enough to be vehicles, *sensu* Dawkins.

We have argued, however, that the framework of social niche construction can remove this impasse. This is for two reasons: first by arguing that more than a contextual role for groups is appropriate in type 1 selection, and second by showing how the cohesive groups that are a prerequisite for type 2 selection can be created. Regarding the first of these, if the group structure can itself change, then just viewing it as part of the static environmental context, in the way that abiotic environmental properties are viewed, misses out a vital part of the story. In particular, viewing groups as a static part of the environment cannot explain how they arose, or how and why the strength of selection pressure they exert on social traits changes. Social niche construction forces a move away from the “groups as context” paradigm, by focussing on groups as themselves being a product of evolution, and not just a background context. This is analogous to the argument made in the niche construction literature for why individual modification of the abiotic environment is different from conventional natural selection, which focusses on how organisms adapt to a static environment (Odling-Smee et al., 2003; Laland and Sterelny, 2006). Like these authors, we argue that it is wrong to contextualise aspects of the environment that can themselves change as a result of individual adaptation.

Moreover, because such “environmental” factors can change adaptively, they themselves require an evolutionary explanation for their current state. In general, we argue that the view of type 1 group selection as individuals simply adapting to the group structure that they find themselves in, is limited in its explanatory power. Of course, social niche construction, and indeed any type of group selection, must result in a change in frequency of certain types of individual, if only because groups are made of individuals. That groups are made of individuals, however, has no bearing on whether their role in evolution is important.

Our second argument is that social niche construction can help to bridge the gap between type 1 and type 2 group selection. This is because it explains how individual adaptations strengthen group selection by creating more cohesive groups. Such greater group cohesion is moving towards the kind of groups that type 2 selection can operate on. Thus, the two seemingly disparate types of group selection are reconciled when one considers that type 1 could evolve into type 2 (see also Okasha 2006 and Michod 2005).

8.2 Evaluation of the methodology and further investigations in the current models

We have advanced our thesis of social niche construction through the development of a logical argument. We have also illustrated this argument through a series of simulation models. When developing the simulation models, we have adopted the philosophy that a model should be no more complex than it needs to be in order to illustrate a particular effect. Thus, we have not added extra factors into the models for the sake of a supposed increase in realism, since the more factors that are added, the harder it is to determine which are necessary for producing any particular effect. Moreover, adding extra factors that we do not have a good understanding of introduces a greater possibility of artefacts being introduced from the particular assumptions made. We have thus aimed to make as few assumptions as possible in our models, and where we do make them, to try to have a principled understanding of what effect they have.

For example, our logical argument did not make reference to diploidy or sexual reproduction, and so we modelled asexual reproduction of haploid genotypes. This is a simpler case to understand, but crucially still allows us to illustrate our argument. Similarly, we considered an aggregation and dispersal metapopulation structure, since this has few parameters and easily elucidates the conditions under which cooperation is selectively advantageous, as discussed in Chapter 3. For the same reasons, this model metapopulation structure has been used by Sober and Wilson (1998) to illustrate the factors affecting the evolution of social behaviour. Our group formation mechanism in Chapter 6 was also chosen so as to satisfy the assumption of our logical argument that individuals with a trait creating a certain population structure tend to find themselves living in that structure more frequently than individuals that do not posses the trait. We satisfied this in a simple manner, by assigning individuals to groups in order of their size preference (Algorithm 6.2). Of course, such global sorting is not biologically realistic, but it should be possible to devise a process that achieves the same result based only on local interactions between organisms. Investigating such a mechanism would make an interesting piece of future work, but was not necessary to illustrate our argument.

Concerning simulation technique, we have not used fully individual-based models. Rather, we have used difference equations on continuous genotype densities to model genotype reproduction. We have thus treated each group as a single freely-mixed population, as is standard in analytical and numerical models of group selection in the literature (indeed, the trait-group concept of group selection defines a group as the largest subset of the population for which interactions *are* freely-mixed), e.g., Wilson (1975a, 1987); Fletcher and Zwick (2004). We have avoided the use of full individual-based models because the extra assumptions that they make can introduce artefacts into the results (Powers and Watson, 2007). For example, in a model where individuals occupy sites on a lattice, one must make an assumption about how the cells making up the lattice are updated: are they updated one at a time, or simultaneously? This innocuous-sounding assumption can in fact have a major impact on the results of the simulation, as was shown by Huberman and Glance (1993). They found that the coexistence of cooperators and selfish individuals in a Prisoner's Dilemma game played on a lattice reported by Nowak and May (1992) only applied when lattice sites were updated synchronously; if the sites were instead updated asynchronously, then selfish behaviour swept to fixation. Because we did not need to use an individual-based model to illustrate our argument, we avoided introducing such extra complications that come purely from the modelling technique. In general, we have tried to use modelling techniques that do not introduce any unnecessary assumptions.

However, now that we have illustrated our logical argument with minimal models, it would be insightful to investigate other factors, one by one, that may affect the results. One such factor that is particularly pertinent is sexual reproduction, for this can reduce the level of linkage disequilibrium when recombination occurs between separate loci (Ridley, 2003). Reduction of linkage disequilibrium between population-structuring traits and socio-behavioural traits would reduce selection for the structuring trait. This is because it would reduce the association between possessing a structuring trait and receiving the benefits of increased cooperation. Thus, the fitness advantage of individuals with the structuring trait that comes from experiencing greater cooperation would be reduced. We should stress, however, that we would not expect recombination to destroy linkage disequilibrium entirely, particularly if individuals mate with others within their structure, since the assortativity on structuring traits that is a premise of our social niche construction argument means that mates will themselves tend to have the same structuring allele. Thus, the degree to which individuals with the same structuring allele tend to live in the same structure will affect the potential of recombination to disrupt linkage disequilibrium. The recombination rate during sexual reproduction will also have an effect (see Ridley 2003 for a discussion of how the recombination rate affects linkage disequilibrium).

As hinted at above, another factor that should have qualitatively similar effects to sexual reproduction is to reduce the probability that individuals end up living in the size of group coded by their size preference allele. In the group formation algorithm presented here (Algorithm 6.2), assigning individuals to groups in order of their size preference, and then creating a new group when adding the next individual would cause group size to exceed the mean preference of its members, means that most individuals end up in the size of group that they prefer. We can relax this, by making group formation based on size preference a stochastic process[2]. In particular, individuals could live in a group that satisfies their size preference with a certain probability. Procedurally, this could be achieved by a modified version of Algorithm 6.2 that added the next individual in the sorted list to the current group with a certain probability, otherwise an individual chosen from a random position in the list would be added instead. Reducing the matching between preference and size of group lived in would reduce the linkage disequilibrium between group size and social behaviour, since a smaller fraction of the individuals living in smaller, more cooperative, groups would also posses the small allele that created those groups. Thus, the benefits of cooperation would fall less preferentially on individuals with the small size allele, thereby reducing the fitness advantage of the allele that comes from social behaviour. Given the discussion of the effects of sexual reproduction above, it would also be insightful to vary both the recombination rate and the degree of preference matching simultaneously.

Our models consider a loose trait-group structure to be initially present in a population. This follows from the notion that all individuals will have some fitness-affecting interactions with others. This applies even at the origin of life, for example, the first replicators are hypothesised to have needed the same chemicals for their reproduction and hence to compete for those resource (Szathmáry, 1999). It is thus not reasonable to think that organisms could ever exist in isolation, for their fitness will always be affected by biotic as well as abiotic factors. At the same time, most organisms are not constantly well-mixed, and so these fitness affects will be localised. Fitness affects from biota therefore create a trait-group structure, as argued by Wilson (1975a, 1980). We thus argue that it is a fundamental mistake to think that a group size of 1, representing an organism in a vacuum, is the natural starting point of a model of the evolution of population structure.

However, the group structure in our model is fairly rigid, in the sense that the groups are discrete and completely isolated between the mixing stage. It would be interesting to examine the evolution of population structure in a less rigid scenario, for example, in an

---

[2] Group formation is stochastic with respect to the socio-behavioural trait in the models.

individual-based model where individuals occupy sites on a lattice and interact with their neighbours. In such a model, the interaction groups would be continuous rather than discrete, and a global mixing stage every t generations could be replaced by a varying dispersal distance of offspring from their parents. We could then examine whether a more defined group structure, more conducive to between-group selection, could evolve. Doing so could determine how well defined the initial population structure needs to be in order for social niche construction to occur. Taken to the extreme, one could even start individuals out in a completely freely-mixed state, and then give individuals a trait that determines how much they adhere to, and hence keep on interacting with, other individuals. This adherence trait could take, for example, the form a chemical secretion, as in extracellular matrix production in bacterial biofilms (Flemming et al., 2007). Although we have argued that such a freely-mixed state is not really a feasible starting point, using such a model would allow us to test the limits of our social niche construction hypothesis. This is because it would allow us to investigate whether any population structure at all is needed at the outset, or if social niche construction can take a completely unstructured population into a highly structured state that supports cooperation. Even if it cannot, it could potentially be bootstrapped by self-organisation (see, for example, Boerlijst and Hogeweg 1991). That is, self-organisation may generate a loose structure that natural selection can act on, and which can provide the starting point for social niche construction.

8.3 Applying the social niche construction framework: directions for future research

Our social niche construction thesis postulates that the benefits of cooperative social behaviour drive the evolution of population structure. This thesis provides an abstract template that can be instantiated by considering the evolution of particular population-structuring and social behavioural traits. We provided one particular instantiation in Chapter 6 by considering the population-structuring trait of initial group size preference, and the socio-behavioural trait of cooperation in public goods production, as modelled by the Snowdrift and Prisoner's Dilemma games. Below, we consider some other possible instantiations of the social niche construction framework, and how they can help to resolve controversial issues in the group selection literature.

8.3.1 Modelling other kinds of social-niche constructing traits and metapopulation structures

One obvious avenue for future research is to explicitly model the evolution of other kinds of social niche-constructing traits. Our modelling work in this dissertation has focussed on the particular trait of initial group size preference which, as discussed in Chapter 7, is particularly relevant to the evolution of a developmental bottleneck that occurs in several of the major transitions. However, our logical argument for social niche construction does not rely on, or even make reference to, this particular trait, or to the particular aggregation and dispersal metapopulation structure that we have modelled. Rather, our argument simply requires the existence of an individual trait that can create a different population structure, and that this different structure provides a greater assortment of social behaviours which leads to increased cooperation being selected for.

When just considering an aggregation and dispersal metapopulation, there are at least two other social niche-constructing traits that can be modelled; these are the number of generations spent within groups before mixing, and the rate of migration that occurs between groups during their growth. Both of these traits can be influenced genetically, in a particular, by a genetic tendency for individuals to disperse from their natal habitat. Moreover, both of the traits directly affect between-group variance in the aggregation and dispersal structure, and hence the assortment of social behaviours. Thus, when evolved concurrently with social behaviour, they can provide another instantiation of our logical argument.

A line of enquiry that appears particularly fruitful is then to consider the concurrent evolution of several social niche-constructing traits simultaneously. For example, both the number of generations spent within groups between mixing episodes, and initial group size preference, could be evolved simultaneously alongside social behaviour. Evolution through more than one dimension of population structure could create the conditions for even greater levels of cooperation to evolve. It is also possible that the selective processes on the separate social niche-constructing traits could interact with each other in a positive feedback process. We see such positive feedback in the evolution of initial group size and cooperation, since a smaller initial group size selects for greater cooperation, and through the generation of linkage disequilibrium this greater cooperation differentially increases the fitness of individuals that create the smaller groups, leading to selection for the small group trait. Thus, the small group trait selects for cooperation, and cooperation selects for the small group trait (see also Breden and Wade 1991). In a similar manner, selection on different social niche-constructing traits, through their effect on social behaviour, could potentially reinforce each other.

Because our argument does not require the aggregation and dispersal metapopulation structure, it is also worth considering social niche construction in other kinds of group structure. One example would be the classical demic structure investigated in the models of the early 1970s (Boorman and Levitt, 1972; Levin and Kilmer, 1974), in which there is no global mixing stage where the groups disband, but instead groups are subject to extinction events with a probability inversely proportional to their frequency of cooperation. When extinction occurs, the now vacant site is recolonised by migrants from other groups, where the number of migrants that a group sends out may be proportional to its frequency of cooperation. Traits such as the deme size and migration rate could again be evolved alongside social behaviour in such a model. An adaptive gradient towards the conditions for greater cooperation can again exist if a small decrease in say, deme size, leads to an increase in cooperation and individual fitness.

Similarly, the concurrent evolution of dispersal traits and social behaviour could be investigated in an explicitly spatial model, such as a viscous population. It has been proven that a purely viscous population, where there is no long-range dispersal but instead offspring remain close to where they were born, does not favour the evolution of cooperation. This is because increased competition within a locality resulting from no dispersal offsets the benefits of helping other individuals in that locality (Taylor, 1992; Wilson et al., 1992; West et al., 2002), as discussed at the beginning of Chapter 3. What has not been addressed in the literature on this topic, however, is whether a population could evolve from a purely viscous state to a structure more conducive to cooperation; for example, a structure in which there is periodic dispersal. Avoiding competition with relatives has long been postulated to influence the evolution of dispersal (Hamilton and May, 1977), but could increased cooperation rather than reduced competition *per se* also be a driver? If an increased dispersal rate induces selection for greater cooperation, then the social niche construction argument in this thesis applies.

In all of these cases, negative frequency-dependent selection on social behaviour should logically have the same effect as in our aggregation and dispersal, initial group size, model. That is, we would expect to see some increased selection for cooperation induced by a small, mutational, change in population structure, from a much larger range of initial conditions. This is because the existence of a stable polymorphism of behaviours means that some between-group variance, or assortment of behaviours, can potentially be generated when the groups are reformed. By contrast, under directional selection, one behaviour tends to be driven extinct over a large range of parameter space, destroying the possibility of any local variation in the frequency of that behaviour. This conclusion should hold for any kind of population structure where local interaction groups are formed, at least partly, by a stochastic sampling of the global population.

8.3.2 Evolution from loose aggregations to highly cohesive groups

One theme in the group selection controversy is a disagreement over how cohesive a group has to be in order for us to say that it is the target of natural selection. Authors such as Wilson and Sober (1994) view fleeting single-generational associations of individuals, as in Wilson's 1975a; 1980 trait-group concept, as the targets of selection. For others, such as Maynard Smith (1987), the groups must have a clear boundary, presumably be multi-generational, and reproduce in a manner such that parent groups beget whole daughter groups similar to themselves. Social niche construction explains how the former can evolve into the latter. For example single-generational trait-groups could evolve, through selection on an individual trait that affected the frequency of global mixing, into much more clearly defined groups that stay together for many organismic generations. Thus, loose trait-group structure could scaffold the evolution of more concrete groups. Specifically, initial trait-group structure may provide the start of an adaptive gradient on a trait that affects the frequency of global mixing, such that a small decrease in the frequency of mixing (from every generation) could select for some increase in cooperation and hence for social niche construction. This follows from the fact that some models (Wilson, 1987), including some of those in this dissertation (Figure 4.2), have shown that spending more than one generation within groups before mixing can increase the effect of group selection (although this is not always the case, for it depends on how the strength of between- and within-group selection changes through the generations; see Figure 3.1). Thus, there could be a selective gradient on moving from 1 to 2 generations between mixing, then 2 to 3, and so on, if each change yields an increase in the level of cooperation that individuals choosing to live in such groups experience.

This process could continue further beyond the evolution of groups that stay together for multiple generations. In particular, the mode of group reproduction could change adaptively. For example groups may initially reproduce only in the loose sense of contributing individuals to a global migrant pool after a number of generations. However, they could evolve to reproduce by sending out a propagule that founds a new group without any intergroup mixing. When this occurs, we can see groups as forming clear parent-offspring lineages, and so we can start to think of groups as “units of evolution” in their own right, *sensu* Maynard Smith (1987). We can also start to think of the decoupling of group and individual fitness, and hence the creation of type 2 group selection. Crucially, the mode of group reproduction is certainly influenced by individual genetic traits, even if it is an emergent product of the interactions between the different traits of the group members. Thus, social niche construction can potentially operate on the mode of group reproduction, in a manner analogous to the way that we have illustrated it operating on group size. In particular, propagule reproduction may increase between-group variance (Wade, 1978) and hence the degree of cooperation that is selected for. Hence, individuals that choose to live in groups founded by propagules could be fitter than those that live in groups founded by aggregation from a global migrant pool, creating an individual adaptive gradient towards propagule reproduction.

The impact of this is that it can explain how new, higher level, units of evolution are created as a result of individual adaptation. That is, individuals evolve to adaptively create groups that satisfy Maynard Smith’s criteria for groups to be units of evolution. We can therefore provide an account of the evolutionary origin of such groups, rather than having to simply posit their existence as a matter of fact. This is essential to providing an adaptive explanation for the major transitions, in which lower-level individuals do create higher-level evolutionary units. Social niche construction provides a general theory for how this occurs, and one that is couched in terms of immediate individual fitness benefit and hence fully compatible with natural selection. More generally, if trait-groups can bootstrap the evolution of more coherent groups, including those that form clear parent-offspring lineages, then it is perhaps less appropriate to sideline their role in evolution as some authors Nunney (1985); Maynard Smith (1987, 1998) seek to. Rather, they can form the foundation of coherent entities that anyone would be happy to call a group.

8.3.3 Weak altruism can potentially bootstrap the evolution of strong altruism through social niche construction

Many authors have sought to demarcate a clear line between strong and weak altruism (Nunney, 1985; West et al., 2007a). In particular, it is often argued that the former can evolve by group selection, while the latter merely evolves by individual selection and self-interest (Nunney, 1985, 1998)[3]. Weakly altruistic traits can, however, support the evolution of group structure, as shown by the models in this dissertation. In particular, they can lead to a group structure that provides a high degree of between-group variance, for example, by leading to a reduction in initial group size. Such a structure could then provide sufficient between-group variance for a *strongly* altruistic trait to evolve. For example, the benefits of weak altruism could drive randomly formed trait-groups to evolve from being single- to multi-generational, thereby providing the positive assortment necessary for strong altruism to evolve (Fletcher and Zwick, 2004). So, although the initial population structure that the organisms lived in may not have supported strong altruism, the evolution of social niche-constructing traits alongside weak altruism could create a structure where strong altruism is then selectively advantageous. Thus, we suggest that through social niche construction, weak altruism can lead to the conditions that favour strong altruism. Rather than viewing weak altruism as not being relevant to group selection, this suggests a role for such social behaviours in creating the kind of population structure in which group selection is highly effective.

8.3.4 Social niche construction in multi-species and cultural settings

Group selection can operate in multi-species communities as well as amongst groups of conspecifics (Wilson, 1992). Similarly, it can operate in situations where the transmission of traits is cultural rather than genetic (Boyd and Richerson, 1990; McElreath et al., 2003). It has been suggested that both of these cases are more conducive to group selection than single-species genetic models have suggested (Boyd and Richerson, 1990; Wilson, 1992). One fundamental reason for this is that in both cases, there is quite likely to be the existence of more than one ESS within a group. In the cultural setting, these may represent different stable social norms, for example as modelled by a coordination game (McElreath et al., 2003). In a multi-species community setting, they may represent

---

[3] As discussed in Chapter 2, this argument does not apply if the groups exist for multiple generations, as the groups considered in this dissertation do.

different species compositions that are attractors of the ecological dynamics, as modelled by the Lotka-Volterra competition equations, for example (Wilson, 1992; Penn, 2003; Penn and Harvey, 2004). As we discussed in Chapter 5, group selection can then perform equilibrium selection (Harsanyi and Selten, 1988) amongst these ESSs, favouring those that are payoff dominant, i.e., increase the mean fitness of the group members.

Social niche construction could operate in both of these settings to increase between-group variance, such that social niche constructors would posses traits which meant that they tended to find themselves in groups that reached a high fitness ESS. One such trait could be a reduction in initial group size, as considered here in the case of single-species genetic traits. Another could be the evolution of associations with particular other species or types, since as discussed in Chapter 7, we would not expect such associations to be vulnerable to cheating if there are multiple ESSs. Hence, the "green-beard" effect (Hamilton, 1964b; Dawkins, 1976) should be more robust in a setting where there are multiple ESSs.

It would also be interesting to examine how the rapid spread of cultural traits, compared to genetic traits, influences social niche construction. For example, if the behavioural trait was cultural, but the population structure-modifying trait genetic, would the fast evolution of behaviour speed up social niche construction or hinder it?

8.4 Concluding remarks

The vast majority of models of social evolution implicitly view the evolution of cooperation as adaptation to population structure. That is, they show how cooperation is individually advantageous given a certain population structure. This vast line of research then focusses on finding ever more population structures under which cooperation can be supported. However, we have argued that such work cannot hope to provide a satisfactory account of the origin of cooperation. This is because these models are silent on the origin of the structures that support cooperation, and so they cannot provide an account of how cooperation has come to be selectively advantageous. If we want to really understand the evolution of cooperation, we need to understand the evolutionary forces that create population structure.

We have shown that population structures can evolve as a result of individual adaptation, because they select for cooperative behaviour and hence increase the fitness of individuals in that structure. Thus, we have shown not only that population structure drives the evolution of cooperation, as in classical theory, but also that cooperation can itself drive the evolution of population structure, a process that we term *social niche construction*. Moreover, we have developed a logical argument that any component of selection pressure on population structure that is due to social behaviour must be in the direction of increased cooperation; that component of selection cannot favour increased selfishness *per se*.

Our social niche construction thesis speaks directly to the group selection controversy. There is a tendency in the social evolution literature to view apparent cases of group selection as being merely frequency-dependent individual selection causing individuals to adapt to the group structure that they find themselves in. On this understanding, groups are treated as part of the context in which ordinary individual selection occurs, that is, they are treated as a static part of an organism's selective environment in the same way that abiotic factors are. We have argued, however, that the origin of cooperation can only be explained by the creation of group structure. Thus, explaining cooperation requires that groups are treated as more than simply part of the selective background, for their creation needs to be explained by evolutionary processes. Moreover, our logical argument suggests that the benefits of cooperation will drive the creation of group structures ever more conducive to group selection. Thus, group selection selects for cooperation, and cooperation selects for more group selection, in a self-reinforcing process. Sidelining the role of group selection in social evolution, by viewing it as merely frequency-dependent individual selection, can overlook this fundamental dynamic.

We have argued that social niche construction is fundamental to providing an adaptive explanation for the major transitions in evolution. Many of the transitions involve not only the evolution of a high degree of cooperation, but also a change in the pre-transitional organisms' population structure that supported such social evolution. Rather than viewing this change as unsystematic, social niche construction provides an individual adaptive explanation for how population structure and cooperation reinforce each other, which in extreme cases can create the prerequisite features for a transition to a new level of individuality, i.e., the suppression of within-group selection and the dominance of between-group selection.

The fact that we have shown that not only does population structure drive the evolution of cooperation, but also that cooperation can drive the evolution of population structure, suggests that cooperation is indeed a fundamental driving force in evolution.

Bibliography

Alexander, R. and Borgia, G. (1978). Group selection, altruism and the levels of organization of life. *Annual Review of Ecology and Systematics*, 9:449–475.

Allee, W. C. (1938). *The Social Life of Animals.* Norton, New York.

Allee, W. C., Emerson, A., Park, O., Park, T., and Schmidt, K. (1949). *Principles of Animal Ecology.* W. B. Saunders, Philadelphia.

Arnold, A. J. and Fristrup, K. (1982). The theory of evolution by natural selection: A hierarchical expansion. *Paleobiology*, 8:113–129.

Avilés, L. (1993). Interdemic selection and the sex ratio: A social spider perspective. *American Naturalist*, 142(2):320–345.

Avilés, L. (1999). Cooperation and non-linear dynamics: An ecological perspective on the evolution of sociality. *Evolutionary Ecology Research*, 1:459–477.

Avilés, L. (2002). Solving the freeloaders paradox: Genetic associations and frequency-dependent selection in the evolution of cooperation among nonrelatives. *Proceedings of the National Academy of Sciences of the USA*, 99(22):14268–14273.

Avilés, L., Abbot, P., and Cutter, A. D. (2002). Population ecology, nonlinear dynamics, and social evolution. I. Associations among nonrelatives. *American Naturalist*, 159(2):115–127.

Axelrod, R. (1984). *The Evolution of Cooperation.* Basic Books, Inc., NY.

Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.

Barrett, M. and Godfrey-Smith, P. (2002). Group selection, pluralism, and the evolution of altruism. *Philosophy and Phenomenological Research*, 65(3):685–691.

Bergstrom, T. C. (2002). Evolution of social behavior: Individual and group selection. *Journal of Economic Perspectives*, 16(2):67–88.

Boerlijst, M. C. and Hogeweg, P. (1991). Spiral wave structure in pre-biotic evolution: Hypercycles stable against parasites. *Physica D: Nonlinear Phenomena*, 48(1):17–28.

Boomsma, J. J. and Franks, N. R. (2006). Social insects: from selfish genes to self organisation and beyond. *Trends in Ecology & Evolution*, 21(6):303–308.

Boorman, S. A. and Levitt, P. R. (1972). Group selection on the boundary of a stable population. *Proceedings of the National Academy of Sciences of the USA*, 69(9):2711–2713.

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

Breden, F. (1990). Partitioning of covariance as a method for studying kin selection. *Trends in Ecology & Evolution*, 5(7):224–228.

Breden, F. and Wade, M. (1991). "Runaway" social evolution: reinforcing selection for inbreeding and altruism. *Journal of Theoretical Biology*, 153(3):323–337.

Breden, F. and Wade, M. J. (1981). Inbreeding and evolution by kin selection. *Ethology and Sociobiology*, 2(1):3–16.

Brown, J. L. (1978). Avian communal breeding systems. *Annual Review of Ecology and Systematics*, 9:123–155.

Bryden, J. and Jansen, V. A. A. (2010). The impact of clonal mixing on the evolution of social behaviour in aphids. *Proceedings of the Royal Society B: Biological Sciences.* Published online before print February 3, 2010.

Buckling, A., Harrison, F., Vos, M., Brockhurst, M. A., Gardner, A., West, S. A., and Griffin, A. (2007). Siderophore-mediated cooperation and virulence in pseudomonas aeruginosa. *FEMS Microbiology Ecology*, 62(2):135–141.

Bull, J. J., Molineux, I. J., and Rice, W. R. (1991). Selection of benevolence in a host-parasite system. *Evolution*, 45(4):875–882.

Bulmer, M. G. and Taylor, P. D. (1980). Sex ratio under the haystack model. *Journal of Theoretical Biology*, 86(1):83–89.

Burmolle, M., Webb, J. S., Rao, D., Hansen, L. H., Sorensen, Soren, J., and Kjelleberg, S. (2006). Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms. *Applied and Environmental Microbiology*, 72(6):3916–3923.

Buss, L. W. (1988). *The Evolution of Individuality.* Princeton University Press.

Canals, J. and Vega-Redondo, F. (1998). Multi-level evolution in population games. *International Journal of Game Theory*, 27(1):21–35.

Charnov, E. L. and Krebs, J. R. (1975). The evolution of alarm calls: Altruism or manipulation? *American Naturalist*, 109(965):107–112.

Craig, D. M. (1982). Group selection versus individual selection: An experimental analysis. *Evolution*, 36(2):271–282.

Crespi, B. J. (2004). Vicious circles: positive feedback in major evolutionary and ecological transitions. *Trends in Ecology & Evolution*, 19:627–633.

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

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* John Murray, London.

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex.* John Murray, London.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31:169–193.

Dawkins, R. (1976). *The Selfish Gene.* Oxford University Press.

Dawkins, R. (1982). *The Extended Phenotype.* Oxford University Press.

Dawkins, R. (1984). Replicators and vehicles. In Brandon, R. N. and Burian, R. M., editors, *Genes, Organisms, Populations: Controversies Over the Units of Selection*, chapter 11, pages 161–180. MIT Press, Cambridge, MA.

Dawkins, R. (1994). Burying the vehicle. *Behavioral and Brain Sciences*, 17(4):616–617.

di Paolo, E. A. (2000). Ecological Symmetry Breaking can Favour the Evolution of Altruism in an Action-response Game . *Journal of Theoretical Biology*, 203(2):135–152.

Doebeli, M. and Hauert, C. (2005). Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology Letters*, 8(7):748–766.

Doebeli, M., Hauert, C., and Killingback, T. (2004). The evolutionary origin of cooperators and defectors. *Science*, 306(5697):859 – 862.

Dugatkin, L. and Reeve, H. (1994). Behavioral ecology and levels of selection: Dissolving the group selection controversy. *Advances in the Study of Behavior*, 23:101–133.

Dugatkin, L. A. (1990). N-person games and the evolution of cooperation: a model based on predator inspection behavior in fish. *Journal of Theoretical Biology*, 142:123–135.

Dugatkin, L. A., Perlin, M., and Atlas, R. (2003). The evolution of group-beneficial traits in the absence of between-group selection. *Journal of Theoretical Biology*, 220:67–74.

Dugatkin, L. A., Perlin, M., and Atlas, R. (2005). Antibiotic resistance and the evolution of group-beneficial traits. ii: A metapopulation model. *Journal of Theoretical Biology*, 236:392–396.

Dugatkin, L. A. and Wilson, D. S. (1993). Fish behaviour, partner choice experiments and cognitive ethology. *Reviews in Fish Biology and Fisheries*, 3(4):368–372.

Dunbar, M. J. (1960). The evolution of stability in marine environments natural selection at the level of the ecosystem. *American Naturalist*, 94(875):129–136.

Emlen, S. T. and Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*, 197:215–223.

Eshel, I. and Cavalli-Sforza, L. L. (1982). Assortment of encounters and evolution of cooperativeness. *Proceedings of the National Academy of Sciences of the USA*, 79(4):1331–1335.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Clarendon, Oxford.

Flack, J. C., Girvan, M., de Waal, F. B. M., and Krakauer, D. C. (2006). Policing stabilizes construction of social niches in primates. *Nature*, 439:426–429.

Flemming, H.-C., Neu, T. R., and Wozniak, D. J. (2007). The EPS matrix: The House of Biofilm Cells. *Journal of Bacteriology*, 189(22):7945–7947.

Fletcher, J. A. and Doebeli, M. (2006). How altruism evolves: assortment and synergy. *Journal of Evolutionary Biology*, 19:1389–1393.

Fletcher, J. A. and Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. *Proceedings of the Royal Society B: Biological Sciences*, 276(1654):13–19.

Fletcher, J. A. and Zwick, M. (2004). Strong altruism can evolve in randomly formed groups. *Journal of Theoretical Biology*, 228:303–313.

Fletcher, J. A. and Zwick, M. (2007). The evolution of altruism: Game theory in multilevel selection and inclusive fitness. *Journal of Theoretical Biology*, 245(1):26–36.

Fletcher, J. A., Zwick, M., Doebeli, M., and Wilson, D. S. (2006). What's wrong with inclusive fitness? *Trends in Ecology and Evolution*, 21(11):597–598.

Foster, K. R., Wenseleers, T., and Ratnieks, F. L. (2006a). Kin selection is the key to altruism. *Trends in Ecology & Evolution*, 21(2):57–60.

Foster, K. R., Wenseleers, T., Ratnieks, F. L., and Queller, D. C. (2006b). There is nothing wrong with inclusive fitness. *Trends in Ecology & Evolution*, 21(11):599–600.

Frank, S. A. (1996). Models of parasite virulence. *Quarterly Review of Biology*, 71:37–78.

Frank, S. A. (1998). *Foundations of Social Evolution.* Princeton University Press, Princeton.

Ghannoum, M. A. and O'Toole, G. A., editors (2004). *Microbial Biofilms*. ASM Press, Washington, DC.

Gildenhuys, P. (2003). The Evolution of Altruism: The Sober/Wilson Model. *Philosophy of Science*, 70(1):27–48.

Gilpin, M. (1975). *Group Selection in Predator-Prey Communities*. Princeton University Press.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford University Press, Oxford.

Goodnight, K. F. (1992). The effect of stochastic variation on kin selection in a budding-viscous population. *American Naturalist*, 140(6):1028–1040.

Gore, J., Youk, H., and van Oudenaarden, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459:253–256.

Grafen, A. (1984). Natural selection, kin selection and group selection. In Krebs, J. and Davies, N., editors, *Behavioural Ecology*, chapter 3, pages 62–84. Blackwell Scientific Publications, Oxford, 2nd edition.

Grafen, A. (1985). A geometric view of relatedness. In Dawkins, R. and Ridley, M., editors, *Oxford Surveys in Evolutionary Biology*, volume 2. Oxford University Press, Oxford.

Grafen, A. (2006). Optimization of inclusive fitness. *Journal of Theoretical Biology*, 238:541–563.

Greig, D. and Travisano, M. (2004). The Prisoner's Dilemma and Polymorphism in Yeast SUC Genes. *Proceedings of the Royal Society B: Biology Letters*, 271(Supplement 3):S25–S26.

Grey, D., Hutson, V., and Szathmary, E. (1995). A re-examination of the stochastic corrector model. *Proceedings of the Royal Society B: Biological Sciences*, 262(1363):29–35.

Griffin, A. S. and West, S. A. (2003). Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. *Science*, 302(5645):634–636.

Griffin, A. S., West, S. A., and Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, 430:1024–1027.

Griffiths, S. W. (2003). Learned recognition of conspecifics by fishes. *Fish and Fisheries*, 4(3):256–268.

Grosberg, R. K. and Strathmann, R. R. (1998). One cell, two cell, red cell, blue cell: the persistence of a unicellular stage in multicellular life histories. *Trends in Ecology & Evolution*, 13(3):112–116.

Haldane, J. B. S. (1932). *The causes of evolution.* Longmans Green, London.

Hamilton, W. D. (1963). The evolution of altruistic behavior. *American Naturalist*, 97(896):354–356.

Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1):1–16.

Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1):17–52.

Hamilton, W. D. (1972). Altruism and related phenomena, mainly in social insects. *Annual Review of Ecology and Systematics*, 3:193–232.

Hamilton, W. D. (1975). Innate social aptitudes in man, an approach from evolutionary genetics. In Fox, R., editor, *Biosocial Anthropology*, pages 133–155. Malaby Press.

Hamilton, W. D. and May, R. M. (1977). Dispersal in stable habitats. *Nature*, 269:578–581.

Hammerstein, P. (2003). *Genetic and Cultural Evolution of Cooperation.* MIT Press.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162:1243–1248.

Harsanyi, J. C. and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games.* MIT Press, Cambridge, MA.

Harvey, P. H., Partridge, L., and Nunney, L. (1985). Evolutionary biology: Group selection and the sex ratio. *Nature*, 313:10–11.

Hauert, C. and Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428:643–646.

Hauert, C., Michor, F., Nowak, M. A., and Doebeli, M. (2006). Synergy and discounting of cooperation in social dilemmas. *Journal of Theoretical Biology*, 239(2):195–202.

Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519.

Hölldobler, B. and Wilson, E. (2009). *The Superorganism: The Beauty, Elegance, and Strageness of Insect Societies.* W.W. Norton.

Huberman, B. and Glance, N. (1993). Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences of the USA*, 90:7716–7718.

Hull, D. (1984). Units of evolution: A metaphysical essay. In Brandon, R. N. and Burian, R. M., editors, *Genes, Organisms, Populations: Controversies Over the Units of Selection*, chapter 10, pages 142–160. MIT Press, Cambridge, MA.

Johnson, M. L. and Gaines, M. S. (1990). Evolution of dispersal: Theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics*, 21:449–480.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56.

Kerr, B. and Godfrey-Smith, P. (2002). Individualist and multi-level perspectives on selection in structured populations. *Biology and Philosophy*, 17:477–517.

Kerr, B., Neuhauser, C., Bohannan, B. J. M., and Dean, A. M. (2006). Local migration promotes competitive restraint in a host-pathogen 'tragedy of the commons'. *Nature*, 442:75–78.

King, N. (2004). The unicellular ancestry of animal development. *Developmental Cell*, 7(3):313–325.

Kitcher, P., Sterelny, K., and Waters, C. K. (1990). The Illusory Riches of Sober's Monism. *The Journal of Philosophy*, 87(3):158–161.

Koenig, W. D. (1981). Reproductive success, group size, and the evolution of cooperative breeding in the acorn woodpecker. *American Naturalist*, 117(4):421–443.

Kokko, H., Johnstone, R. A., and Clutton-Brock, T. H. (2001). The evolution of cooperative breeding through group augmentation. *Proceedings of the Royal Society B: Biological Sciences*, 268(1463):187–196.

Kondrashov, A. S. (1994). Mutation load under vegetative reproduction and cytoplasmic inheritance. *Genetics*, 137(1):311318.

Krause, J. and Ruxton, G. D. (2002). *Living in Groups.* Oxford University Press, Oxford.

Kreft, J.-U. (2004). Biofilms promote altruism. *Microbiology*, 150:2751–2760.

Lachmann, M., Sell, G., and Jablonka, E. (2000). On the advantages of information sharing. *Proceedings of the Royal Society B: Biological sciences*, 267(1450):1287–1293.

Laland, K. N. and Sterelny, K. (2006). Perspective: Seven reasons (not) to neglect niche construction. *Evolution*, 60(9):1751–1762.

Law, R. (1991). The Symbiotic Phenotype: Origins and Evolution. In Margulis, L. and Fester, R., editors, *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*, pages 57–71. MIT Press, Cambridge, MA.

Lehmann, L. and Keller, L. (2006). The evolution of cooperation and altruism - a general framework and a classification of models. *Journal of Evolutionary Biology*, 19(5):1365–1376.

Leigh Jr., E. G. (2009). The group selection controversy. *Journal of Evolutionary Biology*, 23(1):6–19.

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

Levin, S. A., Muller-Landau, H. C., Nathan, R., and Chave, J. (2003). The ecology and evolution of seed dispersal: A theoretical perspective. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):575–604.

Lewontin, R. and Kojima, K. (1960). The evolutionary dynamics of complex polymorphisms. *Evolution*, 14(4):458–472.

Lewontin, R. C. (1970). The units of selection. *Annual Reviews of Ecology and Systematics*, 1:1–18.

Lewontin, R. C. (1982). Organism and environment. In Plotkin, H. C., editor, *Learning, development and culture*, page 151170. Wiley, New York.

Lewontin, R. C. (1983). Gene, organism, and environment. In Bendall, D. S., editor, *Evolution from Molecules to Men*. Cambridge University Press.

Lorenz, K. (1963). *On Aggression*. Harcourt Brace, San Diego.

MacLean, R. C. (2008). The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies. *Heredity*, 100:233–239.

Margulis, L. (1970). *Origin of Eukaryotic Cells: evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth*. Yale University Press.

Matessi, C. and Jayakar, S. D. (1976). Conditions for the evolution of altruism under Darwinian selection. *Theoretical Population Biology*, 9:360–387.

May, R. M., editor (1976). *Theoretical Ecology: Principles and Applications*. Blackwell Scientific Publications, Oxford.

Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, 201:1145–1147.

Maynard Smith, J. (1976). Group selection. *Quarterly Review of Biology*, 51:277–283.

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

Maynard Smith, J. (1987). How to model evolution. In Dupré, J. A., editor, *The Latest on the Best: Essays on Evolution and Optimality*, pages 119–131. MIT Press, Cambridge, MA.

Maynard Smith, J. (1998). The origin of altruism. *Nature*, 393:639–640.

Maynard Smith, J. and Price, G. (1973). The logic of animal conflict. *Nature*, 246:15–18.

Maynard Smith, J. and Szathmáry, E. (1995). *Major Transitions in Evolution*. W. H. Freeman/Spektrum, Oxford.

Mayo, D. G. and Gilinsky, N. L. (1987). Models of group selection. *Philosophy of Science*, 54(4):515–538.

Mayr, E. (1984). The unity of genotype. In Brandon, R. N. and Burian, R. M., editors, *Genes, Organisms, Populations: Controversies Over the Units of Selection*, chapter 6, pages 69–84. MIT Press, Cambridge, MA.

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

McElreath, R., Boyd, R., and Richerson, P. J. (2003). Shared norms can lead to the evolution of ethnic markers. *Current Anthropology*, 44:122130.

Michod, R. E. (1982). The theory of kin selection. *Annual Review of Ecology and Systematics*, 13:23–55.

Michod, R. E. (1983). Evolution of the first replicators. *American Zoology*, 23:5–14.

Michod, R. E. (1997). Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. *American Naturalist*, 149(4):607–645.

Michod, R. E. (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton University Press.

Michod, R. E. (2005). On the transfer of fitness from the cell to the multicellular organism. *Biology and Philosophy*, 20(5):931–1050.

Michod, R. E. and Roze, D. (2000). Some aspects of reproductive mode and the origin of multicellularity. *Selection*, 1:97–109.

Michod, R. E. and Sanderson, M. (1985). Behavioural structure and the evolution of social behavior. In Greenwood, P. and Slatkin, M., editors, *Evolution - Essays in Honour of John Maynard Smith*, chapter 7. Cambridge University Press.

Mills, R. and Watson, R. A. (2009). Symbiosis enables the evolution of rare complexes in structured environments. In *Proceedings of the 10th European Conference on Artificial Life*. To appear.

Mitteldorf, J. and Wilson, D. S. (2000). Population viscosity and the evolution of altruism. *Journal of Theoretical Biology*, 204:481–496.

Muir, W. (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science*, 75(4):447–458.

Nowak, M. A. and May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, 359:826–829.

Nowak, M. A. and May, R. M. (1993). The spatial dilemmas of evolution. *International Journal of Bifurcation and Chaos*, 3(1):35–78.

Nunney, L. (1985). Group selection, altruism, and structured-deme models. *American Naturalist*, 126(2):212–230.

Nunney, L. (1998). Are we selfish, are we nice, or are we nice because we are selfish. *Science*, 281(5383):1619 – 1621.

Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2003). *Niche construction: the neglected process in evolution.* Monographs in population biology; no. 37. Princeton University Press.

Odum, H. T. and Allee, W. C. (1954). A note on the stable point of populations showing both intraspecific cooperation and disoperation. *Ecology*, 35(1):95–97.

Oechssler, J. (1997). Decentralization and the coordination problem. *Journal of Economic Behavior & Organization*, 32(1):119–135.

Okasha, S. (2001). Why won't the group selection controversy go away? *The British Journal for the Philosophy of Science*, 52(1):25–50.

Okasha, S. (2002). Genetic relatedness and the evolution of altruism. *Philosophy of Science*, 69(1):138–149.

Okasha, S. (2005). Maynard Smith on the levels of selection question. *Biology and Philosophy*, 20(5):989–1010.

Okasha, S. (2006). *Evolution and the Levels of Selection.* Clarendon Press.

Olivieri, I., Michalakis, Y., and Gouyon, P.-H. (1995). Metapopulation genetics and the evolution of dispersal. *American Naturalist*, 146(2):202–228.

Orians, G. H. (1969). On the evolution of mating systems in birds and mammals. *American Naturalist*, 103(934):589–603.

Pacheco, J. M., Santos, F. C., Souza, M. O., and Skyrms, B. (2009). Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proceedings of the Royal Society B*, 276(1655):315–321.

Packer, C. and Ruttan, L. (1988). The evolution of cooperative hunting. *American Naturalist*, 132(2):159–198.

Pál, C. and Papp, B. (2000). Selfish cells threaten multicellular life. *Trends in Ecology & Evolution*, 15(9):351–352.

Pál, C. and Szathmáry, E. (2000). The concept of fitness and individuality revisited. *Journal of Evolutionary Biology*, 13:348–357.

Peck, J. R. (1992). Group selection, individual selection, and the evolution of genetic drift. *Journal of Theoretical Biology*, 159(2):163–187.

Peck, J. R. (2004). Sex causes altruism. Altruism causes sex. Maybe. *Proceedings of the Royal Society B: Biological Sciences*, 271:993–1000.

Peck, J. R. and Feldman, M. W. (1988). Kin selection and the evolution of monogamy. *Science*, 240(4859):1672–1674.

Penn, A. S. (2003). Modelling artificial ecosystem selection: A preliminary investigation. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J. T., and Ziegler, J., editors, *Advances in Artificial Life: Proceedings of the 7th European Conference on Artificial Life (ECAL 2003)*, volume 2801/2003 of *Lecture Notes in Computer Science*, pages 659–666. Springer Berlin / Heidelberg.

Penn, A. S. (2006). *Ecosystem Selection: Simulation, Experiment and Theory.* PhD thesis, University of Sussex.

Penn, A. S. and Harvey, I. (2004). The role of non-genetic change in the heritability, variation and response to selection of artificially selected ecosystems. In *Proceedings of the Ninth International Conference on Artificial Life*, pages 352–357. MIT Press.

Pepper, J. W. (2000). Relatedness in trait group models of social evolution. *Journal of Theoretical Biology*, 206(3):355–368.

Pepper, J. W. (2007). Simple models of assortment through environmental feedback. *Artificial Life*, 13:1–9.

Pfeiffer, T. and Bonhoeffer, S. (2003). An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences of the USA*, 100(3):1095–1098.

Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and competition in the evolution of ATP-producing pathways. *Science*, 292(5516):504–507.

Pollock, G. B. (1983). Population viscosity and kin selection. *American Naturalist*, 122(6):817–829.

Powers, S. T., Mills, R., Penn, A. S., and Watson, R. A. (2009). Social environment construction provides an adaptive explanation for new levels of individuality. In *Proceedings of the ECAL 2009 Workshop on Levels of Selection and Individuality in Evolution: Conceptual Issues and the Role of Artificial Life Models.*

Powers, S. T., Penn, A. S., and Watson, R. A. (2007). Individual Selection for Cooperative Group Formation. In Almeida e Costa, F., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life*, volume 4648/2007 of *Lecture Notes in Computer Science*, pages 585–594. Springer Berlin / Heidelberg.

Powers, S. T., Penn, A. S., and Watson, R. A. (2008). The efficacy of group selection is increased by coexistence dynamics within groups. In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 498–505, Cambridge, MA. MIT Press.

Powers, S. T. and Watson, R. A. (2007). Investigating the evolution of cooperative behaviour in a minimally spatial model. In Almeida e Costa, F., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life*, volume 4648/2007 of *Lecture Notes in Computer Science*, pages 605–614. Springer Berlin / Heidelberg.

Powers, S. T. and Watson, R. A. (2008). The group selection debate and ALife: Weak altruism, strong altruism, and inclusive fitness (abstract). In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, page 796, Cambridge, MA. MIT Press.

Powers, S. T. and Watson, R. A. (2009). Evolution of individual group size preferences can increase group-level selection and cooperation. In Kampis, G., Szathmáry, E., Fernando, C., Jelasity, M., Jordán, F., orincz, A. L., and Scheuring, I., editors, *Advances in Artificial Life: Proceedings of the 10th European Conference on Artificial Life*, volume 5777, 5778 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg.

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

Pruitt, J. N. and Riechert, S. E. (2009). Frequency-dependent success of cheaters during foraging bouts might limit their spread within colonies of a socially polymorphic spider. *Evolution*, 63(11):2966 – 2973.

Queller, D. and Goodnight, K. (1989). Estimating relatedness using genetic markers. *Evolution*, 43:258–275.

Queller, D. C. (1985). Kinship, reciprocity and synergism in the evolution of social behaviour. *Nature*, 318:366–367.

Queller, D. C. (1992a). Does population viscosity promote kin selection? *Trends in Ecology & Evolution*, 7(10):322–324.

Queller, D. C. (1992b). Quantitative genetics, inclusive fitness and group selection. *American Naturalist*, 139(3):540–558.

Queller, D. C. (1994). Genetic relatedness in viscous populations. *Evolutionary Ecology*, 8:70–73.

Queller, D. C. (1997). Cooperators since life began. *The Quarterly Review of Biology*, 72(2):184–188.

Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1403):1647–1655.

Queller, D. C. (2004). Kinship is relative. *Nature*, 430(6):975–976.

Rankin, D. J., Bargum, K., and Kokko, H. (2007). The tragedy of the commons in evolutionary biology. *Trends in Ecology & Evolution*, 22(12):643–651.

Reeve, H. K. and Keller, L. (1999). Levels of selection: Burying the units-of-selection debate and unearthing the crucial new issues. In Keller, L., editor, *Levels of Selection in Evolution*, chapter 1. Princeton University Press.

Richerson, P. J. and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution.* University of Chicago Press, Chicago and London.

Ridley, M. (2003). *Evolution.* John Wiley & Sons, 3rd edition.

Rodman, P. S. (1981). Inclusive fitness and group size with a reconsideration of group sizes in lions and wolves. *American Naturalist*, 118(2):275–283.

Ross-Gillespie, A., Gardner, A., West, S. A., and Griffin, A. S. (2007). Frequency dependence and cooperation: Theory and a test with bacteria. *American Naturalist*, 170(3):331–342.

Roze, D. and Michod, R. E. (2001). Mutation, multilevel selection, and the evolution of propagule size during the origin of multicellularity. *American Naturalist*, 158(6):638–654.

Santos, F. C., Pacheco, J. M., and Lenaerts, T. (2006). Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences of the USA*, 103:3490–3494.

Santos, M. and Szathmáry, E. (2008). Genetic hitchhiking can promote the initial spread of strong altruism. *BMC Evolutionary Biology*, 8(281).

Seeley, T. D. (1997). Honey bee colonies are group-level adaptive units. *American Naturalist*, 150:S22–S41.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society B*, 13:238–241.

Skyrms, B. (2002). Critical commentary on Unto Others. *Philosophy and Phenomenological Research*, 65(3):697–701.

Skyrms, B. (2004). *The Stag Hunt and Evolution of Social Structure.* Cambridge University Press, Cambridge.

Slobodkin, L. B. (1964). Experimental populations of hydrida. *Journal of Animal Ecology*, 52(supplement):131–148.

Snowdon, J. R., Powers, S. T., and Watson, R. A. (2009). Moderate contact between subpopulations promotes evolved assortativity enabling group selection. In *Proceedings of the 10th European Conference on Artificial Life.*

Sober, E. (1984). *The Nature of Selection: Evolutionary Theory in Philosophical Focus.* Bradford/MIT Press.

Sober, E. (1987). Comments on Maynard Smith's"how to model evolution". In Dupré, J., editor, *The Latest on the Best: Essays on Evolution and Optimality*, pages 133–45, Cambridge, MA. MIT Press.

Sober, E. (1990). The Poverty of Pluralism: A Reply to Sterelny and Kitcher. *The Journal of Philosophy*, 87(3):151–158.

Sober, E. (1992). The evolution of altruism: Correlation, cost, and benefit. *Biology and Philosophy*, 7(2):177–187.

Sober, E. and Lewontin, R. C. (1982). Artifact, cause and genic selection. *Philosophy of Science*, 49(2):157–180.

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

Sterelny, K. (1996). The Return of the Group. *Philosophy of Science*, 63(4):562–584.

Sterelny, K. and Kitcher, P. (1988). The return of the gene. *The Journal of Philosophy*, 85(7):339–361.

Sugden, R. (1986). *The Economics of Rights, Cooperation and Welfare.* Blackwell, Oxford and New York.

Számadó, S., Szalaib, F., and Scheuring, I. (2008). The effect of dispersal and neighbourhood in games of cooperation. *Journal of Theoretical Biology*, 253(2):221–227.

Szamado, S. and Szathmary, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution*, 21(10):555–561.

Szathmáry, E. (1999). The first replicators. In Keller, L., editor, *Levels of Selection in Evolution*, chapter 3. Princeton University Press.

Szathmáry, E. and Demeter, L. (1987). Group selection of early replicators and the origin of life. *Journal of Theoretical Biology*, 128(4):463–486.

Taylor, P. D. (1992). Altruism in viscous populations an inclusive fitness model. *Evolutionary Ecology*, 6(4):352–356.

Taylor, P. D. and Frank, S. A. (1996). How to make a kin selection model. *Journal of Theoretical Biology*, 180:27–37.

Taylor, P. D. and Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.

Taylor, P. D. and Wilson, D. S. (1988). A mathematical model for altruism in haystacks. *Evolution*, 42(1):193–196.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35–57.

Uyenoyama, M. and Feldman, M. W. (1980). Theories of kin and group selection: A population genetics perspective. *Theoretical Population Biology*, 17(3):380–414.

Van Huyck, J., Battalio, R., and Beil, R. (1990). Tacit coordination games, strategic uncertainty and coordination failure. *American Economic Review*, 80:234–248.

van Veelen, M. (2009). Group selection, kin selection, altruism and cooperation: When inclusive fitness is right and when it can be wrong. *Journal of Theoretical Biology*, 259(3):589–600.

Varma, A. and Chincholkar, S., editors (2007). *Microbial Siderophores*, volume 12 of *Soil Biology*. Springer, Berlin / Heidelberg.

Wade, M. J. (1976). Group selections among laboratory populations of tribolium. *Proceedings of the National Academy of Sciences of the USA*, 73(12):4604–4607.

Wade, M. J. (1977). An experimental study of group selection. *Evolution*, 31(1):134–153.

Wade, M. J. (1978). A critical review of the models of group selection. *The Quarterly Review of Biology*, 53:101–114.

Wade, M. J. (1980). Kin selection: Its components. *Science*, 210(4470):665 – 667.

Wade, M. J. and Breden, F. (1981). Effect of inbreeding on the evolution of altruistic behavior by kin selection. *Evolution*, 35(5):844–858.

Watnick, P. and Kolter, R. (2000). Biofilm, City of Microbes. *Journal of Bacteriology*, 182(10):2675–2679.

Watson, R. A., Palmius, N., Mills, R., Powers, S. T., and Penn, A. S. (2009). Can selfish symbioses effect higher-level selection? In Kampis, G., Szathmáry, E., Fernando, C., Jelasity, M., Jordán, F., orincz, A. L., and Scheuring, I., editors, *Advances in Artificial Life: Proceedings of the 10th European Conference on Artificial Life*, volume 5777, 5778 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg.

Weibull, J. W. and Salomonsson, M. (2006). Natural selection and social preferences. *Journal of Theoretical Biology*, 239(1).

West, S., Griffin, A., and Gardner, A. (2007a). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20:415–432.

West, S. A. and Buckling, A. (2003). Cooperation, virulence and siderophore production in bacterial parasites. *Proceedings of the Royal Society B: Biological Sciences*, 270(1510):37–44.

West, S. A., Griffin, A. S., and Gardner, A. (2007b). Evolutionary explanations for cooperation. *Current Biology*, 17(16):R661–R672.

West, S. A., Griffin, A. S., and Gardner, A. (2008). Social semantics: how useful has group selection been? *Journal of Evolutionary Biology*, 21(1):374–385.

West, S. A., Pen, I., and Griffin, A. S. (2002). Cooperation and competition between relatives. *Science*, 296(5565):72 – 75.

Wheeler, W. M. (1911). The ant-colony as an organism. *Journal of Morphology*, 22(2):307 – 325.

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

Wilkinson, G. S. (1984). Reciprocal food-sharing in the vampire bat. *Nature*, 308:181–184.

Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton University Press.

Williams, G. C. and Nesse, R. M. (1991). The dawn of darwinian medicine. *The Quarterly Review of Biology*, 66(1):1–22.

Williams, G. C. and Williams, D. C. (1957). Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution*, 11:32–39.

Wilson, D. S. (1975a). A theory of group selection. *Proceedings of the National Academy of Sciences of the USA*, 72(1):143–146.

Wilson, D. S. (1977). Structured demes and the evolution of group-advantageous traits. *American Naturalist*, 111(977):157–185.

Wilson, D. S. (1979). Structured demes and trait-group variation. *American Naturalist*, 113(4):606–610.

Wilson, D. S. (1980). *The Natural Selection of Populations and Communities.* Benjamin/Cummings, California.

Wilson, D. S. (1983a). The effect of population structure on the evolution of mutualism: A field test involving burying beetles and their phoretic mites. *American Naturalist*, 121(6):851–870.

Wilson, D. S. (1983b). The group selection controversy: History and current status. *Annual Review of Ecology and Systematics*, 14:159–187.

Wilson, D. S. (1987). Altruism in mendelian populations derived from sibling groups: The haystack model revisited. *Evolution*, 41(5):1059–1070.

Wilson, D. S. (1990). Weak altruism, strong group selection. *Oikos*, 59(1):135–140.

Wilson, D. S. (1992). Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. *Ecology*, 73(6):1984–2000.

Wilson, D. S. (1997). Altruism and organism: Disentangling the themes of multilevel selection theory. *American Naturalist*, 150: Supplement:S122–S134.

Wilson, D. S. (2004). What is wrong with absolute individual fitness? *Trends in Ecology and Evolution*, 19(5):245–248.

Wilson, D. S. (2007). Social semantics: toward a genuine pluralism in the study of social behaviour. *Journal of Evolutionary Biology*, 21(1):368–373.

Wilson, D. S. and Colwell, R. K. (1981). Evolution of sex ratio in structured demes. *Evolution*, 35(5):882–897.

Wilson, D. S. and Dugatkin, L. A. (1997). Group selection and assortative interactions. *The American Naturalist*, 149(2):336–351.

Wilson, D. S., Pollock, G. B., and Dugatkin, L. A. (1992). Can altruism evolve in purely viscous populations? *Evolutionary Ecology*, 6(4):331–341.

Wilson, D. S. and Sober, E. (1989). Reviving the superorganism. *Journal of Theoretical Biology*, 136(3):337–356.

Wilson, D. S. and Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17(4):585–654.

Wilson, D. S. and Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology*, 82(4):327–348.

Wilson, E. O. (1975b). *Sociobiology: The New Synthesis.* Harvard University Press.

Wolpert, L. and Szathmáry, E. (2002). Multicellularity: Evolution and the egg. *Nature*, 420:745.

Wright, S. (1945). Tempo and mode in evolution: A critical review. *Ecology*, 26:415–419.

Wynne-Edwards, V. C. (1962). *Animal Dispersion in Relation to Social Behaviour*. Oliver and Boyd.

Zea-Cabrera, E., Iwasa, Y., Levin, S., and Rodrguez-Iturbe, I. (2006). Tragedy of the commons in plant water use. *Water Resources Research*, 42(6).