A Causal Interpretation of Selection Theory

By

Peter Avery Gildenhuys

BA, University of Western Ontario, 1998

MA, Northwestern University, 2001

Submitted to the Graduate Faculty of

The School of Arts and Sciences in partial fulfillment

of the requirements for the degree of Doctor of Philosophy

University of Pittsburgh

2009

University of Pittsburgh

School of Arts and Sciences

This dissertation was presented

by

Peter Avery Gildenhuys.

It was defended on

August 06, 2009

and approved by

Sandra Mitchell, Professor, Department of History and Philosophy of Science

Kenneth Schaffner, University Professor, Department of History and Philosophy of Science

Edouard Machery, Associate Professor, Department of History and Philosophy of Science

Robert Brandom, Distinguished Professor of Philosophy, Department of Philosophy

Dissertation Director: James Lennox, Professor, Department of History and Philosophy of Science

A CAUSAL INTERPRETATION OF SELECTION THEORY

Peter Gildenhuys, PhD

University of Pittsburgh, 2009

The following dissertation is an inferentialist account of classical population genetics. I present the theory as a definite body of interconnected inferential rules for generating mathematical models of population dynamics. To state those rules, I use the notion of causation as a primitive. First, I put forward a rule stating the circumstances of application of the theory, one that uses causal language to pick out the types of entities over which the theory may be deployed. Next, I offer a rule for grouping such entities into populations based on their competitive causal relationships. Then I offer a general algorithm for generating classical population genetics models for such populations on the basis of what causal influences operate within them.

Dynamical models in population genetics are designed to demystify natural phenomena, chiefly to show how adaptation, altruism, and genetic polymorphism can be explained in terms of natural rather than supernatural processes. In order for the theory to serve this purpose, it must be possible to understand, in a principled fashion, when and how to deploy the theory. By presenting the theory as a system of ordered inferential rules that takes causal information as its critical input and yields dynamical models as its outputs, I show explicitly how classical population genetics functions as a non-circular theoretical apparatus for generating explanations. The generalization of the theory achieved by presenting it using causal vocabulary shows how the scope of the theory of natural selection extends beyond its traditional domain of systems of individuals distinguished by genetic variations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF EQUATIONS

1.0 SELECTION THEORY

My subject is selection theory, the study of natural selection. This opening section serves simply to set the boundaries of that term, as well as is possible, with respect to nearby vocabulary that is more widely deployed.

What I call selection theory would be treated by most people as evolutionary theory, or at least a part of evolutionary theory. But the term "evolutionary theory" is indeterminate and may include subject matter that falls outside the purview of my concern. For instance, developmental systems theorists explicitly endorse an integrated approach to the study of evolution and development, one designed to explain far more than just the spread of variations within populations (Griffiths and Gray 2001, 196). Such an integrated approach is not undertaken here. Others understand evolutionary theory as nothing more than a diffuse collection of models, principles, generalizations, and explanations. One recent writer listed Wright's shifting balance theory, Fisher's fundamental theorem of natural selection, Eldredge and Gould's theory of punctuated equilibria, and the thermoregulatory theory of the origin of feathers to get across the content of evolutionary theory (Shanahan 2003, 164). So while *evolutionary theory* is many things to many people, selection theory is a term that is not so widespread as to invite misunderstanding.

Besides the indeterminate reference of "evolutionary theory," there are a couple of additional reasons to reject that label for my subject matter: natural selection can explain a lack of evolution, as in frequency-dependent selection for stable polymorphisms, and some

evolutionary change may be inexplicable on the basis of selection theory, such as that prompted not by competition between variants but by global warming, which has certainly caused evolutionary change within some taxa. It would be a mistake to put the word "evolutionary" into the name of a theory that both explains more than just evolution and also fails to explain some of it.

Population genetics is the theory that probably comes closest to picking out what I intend to mean by the term "selection theory," but population genetics, as it is usually presented, is too closely connected with specifically genetic variations to fully serve my purposes. Officially, my view is that chromatin marking schemes, cell membranes, ideas, even whole organisms, can be treated in the way that alleles are treated in classical population genetics models, at least sometimes. So while alleles are the paradigm entities over which selection theory is deployed, its use is not restricted to them. In Chapter 3 I put forth the term "competitor" as a generalization of "allele"; in that chapter I take an explicit and rigid stance on the question of what sorts of systems can be understood using selection theory. For now, what I am calling selection theory should be understood as a generalization of population genetics. It is a theory that provides the means for inferring the dynamics of a subset of natural systems, ones containing competitors, where competitors include at least alleles.

1.1 EXPLANATION AND DEMYSTIFICATION

Historically, Darwin, the founder of selection theory, sought to demonstrate that evolution and adaptation were explicable *using* his new theory of natural selection (Lennox 1991). But the imaginary scenarios that Darwin deploys in the *Origin* to show how otherwise mysterious phenomena are explicable in terms of evolution by natural selection are not explanations of those

phenomena. Indeed, they are not even putative explanations: Darwin's imaginary scenarios are not even supposed to capture how the evolution of real populations actually occurred. Nonetheless Darwinian demonstrations that biological structures are explicable in terms of natural selection remain legitimate argumentative tools against the stance that the same structures *could not* have evolved by natural selection. Such demonstrations show that otherwise puzzling and designed-appearing structures could indeed have evolved by natural selection, and these form part of a broader assault by Darwin on the creationist invocation of the supernatural to explain features of the biological world. We do not need supernatural explanations of adaptations to make sense of their designed-appearing character.

While Gould and Lewontin's (1979) are right to attack uses of imaginary scenarios in which they are presented as full-fledged explanations, Darwin's imaginary illustrations find a role in a more modest use of the theory of natural selection: they show how an adaptation is explicable by natural selection, how it is possible for such adaptations to exist, and so act as counters to the impossibility arguments, launched by Darwin's opponents, to the effect that organs of extreme perfection, such as the eye, could not have evolved by natural selection (Lennox 1991; see discussion in Brandon 1990, 176-184).

Contemporary selection theory provides a means for inferring the dynamics of biological systems, but the point of having a theory that makes this possible is to have something that demystifies biological phenomena. I pick out three biological phenomena that it is the point of contemporary selection theory to demystify: adaptation, genetic polymorphism, and altruism. The presence of designed-appearing structures in the natural world is the chief mystery that selection theory is used to solve, but genetic polymorphism and altruism are mysterious phenomena too, though probably not quite as mysterious. Selection theory makes sense of the mysterious phenomena just listed by showing how they *could be* explained.

Using a theory to explain phenomena and using it to show how they are explicable are not the same things. However, both tasks can only be effected using with a theory that has an explanatory structure. So the main thrust of this work is to offer an explicit account of the workings of selection theory that demonstrates how the theory can function as an *explanatory* theory.

While the most straightforward way to argue that selection theory must have an explanatory structure would be to argue that it explains, I do not argue this way, even though I believe that selection theory can be, and indeed is, used to explain things. For instance, we have very good evidence that spatially and temporally variable selection is responsible for the persistence of a color polymorphism in *Linanthus parryae* (Schemske and Bierzychudek 2001; Turelli, Schemske, and Bierzychudek 2001; Schemske and Bierzychudek 2007). Instead, I argue 1) that the chief (though not sole) role of selection theory is to demystify adaptations, genetic polymorphism, and altruism, but not necessarily to explain them; 2) that a theory that demystifies must have an explanatory structure, for it must explain the dynamics of plausible but imaginary populations, and hence 3) that selection theory must have an explanatory structure. I argue in this

way because I regard deployments of the theory over imaginary populations, ones that do not even *purport* to explain anything as legitimate, indeed seminal, deployments of the theory.

Actual explanatory deployments of selection theory over real systems take a backseat to imaginary ones not only historically, but in other ways too. Actual explanatory uses of selection theory are much more dispensable, for instance. Were it to turn out that many actual deployments of selection theory were misguided, we would be left without an explanation of the target structures, but we would still be left with a theory that showed how adaptations could evolve. Even if selection theory could never be used to actually explain anything owing, say, to insuperable epistemic limitations, it would still have its demystification function, and it would have to have an explanatory structure to fulfill that function. Were the theory to cease to fulfill its function in making mysterious biological structure explicable, however, we would be left with a major mystery on our hands, for we would lack any account of how designed-appearing biological structures even *could* have come to be.

While I think it is interesting and important that the main point of selection theory is demystification, and it is because selection theory must demystify that it must have an explanatory structure, I nonetheless propose to move on and simply show that my stance is coherent not only with evolutionary theory as Darwin understood it, but also with research in contemporary population genetics. Someone who would contend with me over this issue would almost certainly argue that the role of selection theory is not to make organic phenomena explicable, but instead to explain them. Someone who takes that view effectively grants me the claim that I ultimately want to secure, namely that selection theory must have an explanatory structure, but undertakes the claim for different reasons. So we can simply agree to disagree about the justification, and agree that selection theory must have an explanatory structure.

1.2 EXPLICABILITY AND POPULATION GENETICS

The view that the main point of selection theory is about demonstrating the explicability of biological phenomena jibes nicely with population genetics practice (Plutynski 2004). Crucial to the project of using population genetics to demystify, or to provide of "how possibly" explanations, is the drawing of analytic results from population genetics models concerning the dynamics of the systems to which the models apply. Historically, the main research activity of population geneticists has been to develop models from which analytic results concerning the dynamics of populations of alleles can be drawn. These models are almost always concerned with the long-term dynamics of populations (Gale 1990, 40).

Consider, for example, Fisher's "fundamental theorem of natural selection." In one interpretation of this principle, Fisher is taken to have shown that the rate of increase in mean fitness in a population is equal to the additive component of the genetic variance in fitness, at least for a subset of evolutionary situations. Ewens argues that the fundamental theorem of Fisher is deeper and more general than this (2004, 18, 64-67), but matters of historical interpretation aside, my point remains the same: Fisher contributed to the program of demystifying adaptation by showing how, for a certain range of cases, selection leads to increased fitness, thereby providing a rigorous mathematical formulation of how selection promotes the spread of more adaptive structures (This is true, provided mutant alleles that produce incrementally superior variations in such structures recur in the population, something that can be ascertained independently.)

From the same perspective, consider population genetics models that show how genetic variation is maintained in populations by selection. Genetic polymorphism has been a major

explanatory target among population geneticists since the 1940s (Levins 2004, 22); Gillespie begins the preface to his treatise challenging the neutral theory, *The Causes of Molecular Evolution*, with the claim, "naturally occurring genetic variation is an enigma" (Gillespie 1991, vii) . Models of overdominance, in which the relative fitness of the heterozygote is greater than that of either homozygotes, are the best known models that show how genetic variation can persist in a population owing to selection, but the persistence of genetic variation is also an outcome, too, of variable selection models, sex-dependent selection models, frequency-dependent selection models, and many others. Models demonstrating how such variation could persist in population owing to a variety of different sorts of selection regimes contribute to the project of demystifying genetic polymorphism. Often, no effort is made to apply polymorphism-yielding models to actual populations; many variable selection models, for instance, are far too complicated for this. But such models nonetheless contribute to demystifying the extent of genetic polymorphism by showing how it is explicable as the result of selection.

As another illustration of how population genetics is used to demystify genetic polymorphism, consider how the neutral theory, the rival contender to selectionist accounts of genetic polymorphism, explains genetic polymorphism as the result of random drift.[1] The debate between balancing selectionists and neutralists over how much genetic polymorphism can be accounted for by each camp is ongoing. But they are competing for the same territory, population geneticists in both camps seek to account for polymorphism by offering models that yield persistent polymorphisms as their inferential outputs.

As one last example of this sort of thing, consider the mystery of biological altruism. At first glance, we should expect very little, indeed no, instances of behaviors in the natural world that benefit others at the expense of the individual performing them, for we would expect that

[1] See chapter 7 for a specification of the causal notion of drift.

selfish individuals would do not act in such a fashion would invade populations of altruists and spread at the expense of their counterparts until no altruists remained. But altruism exists in nature. Accordingly, population geneticists have been in the business of demystifying altruism too, generating specialized models in which altruism spreads, or is at least maintained, in populations despite its having some negative consequences for the relative fitness of the altruists (e.g., Hamilton 1964; Sober and Wilson 1998; Michod 1999; Dessalles 2007). Sober and Wilson note that even the most famous case of altruism in nature, that of the brainworm parasite *D. dentriticum*, is not one over which population genetics models have actually been deployed (Sober and Wilson 1998, 30). All that has been shown is that it is the sort of system in which altruism could be explained by (a class of) formal population genetics models. Population geneticists settle for demystifying altruism.

How do models for which mean fitness increase, the maintenance of polymorphism, and the persistence of altruism are the *explananda* serve to discharge the mysteries generated by the biological phenomena in the natural world that they target given that such models are rarely if ever applied to actual populations? The models are rarely even tested on natural populations (an interesting fact in its own right). So such models do not explain by showing how populations do indeed evolve in accordance with what the models predict. Furthermore, the models from which analytic results can be drawn are necessarily simple. Many selection processes in nature will be much more complex than those that can be captured by a model from which analytic results can be drawn. Multilocus selection is the standard example of a complicated but potentially widespread phenomenon that is analytically intractable.

Models of mean fitness increase, persistent polymorphism, and altruism do not explain so much as they demystify: they contribute to showing how adaptation, genetic polymorphism, and

altruism are explicable using the formal techniques of population genetics. Phenomena such as biological altruism are alarming until it is shown, by means of a variety of models, that biological altruism can be maintained in populations. Researchers do not go on to try to fit every last altruistic behavior to this or that evolutionary model; rather, the idea is that those to whom the models are presented will no longer find the various instances of altruistic behavior in the biological world so mysterious.

Because the central function of selection theory, both as used by Darwin and as used in quantitative form by his intellectual descendants, is to demystify rather than to explain, I do not go so far as to claim that selection theory explains adaptation. But nonetheless, I claim that the theory must be structured as an explanatory theory if it is to fulfill its function as a demystification device. This claim is true not only of the theory in its qualitative form, as Darwin initially conceived; it is equally true of the quantitative formulations of population genetics theory. Demonstrating that biological traits are explicable requires one to have an explanatory theory, the sort of theory that could explain them. That polymorphism and altruism are explicable in terms of selection theory is demonstrated by exhibiting these as the inferential outputs of a formal mathematical theory with an explanatory structure.

That selection theory must be shown to have an explanatory structure places some constraints on what an account of selection theory must be like. In the next three sections of this chapter, I explore three of those constraints.

1.3 SELECTION THEORY AS A SET OF INFERENCE RULES

A theory with an explanatory structure must be one with an inferential structure. Not all inferences are explanatory, but all explanations are inferences of some sort or another. I do not

motivate this constraint on explanation, but it is quite a weak constraint, and I expect my reader to grant it. It is very difficult to imagine a putative *explanans* not inferentially related to its *explanandum* nonetheless being counted as an explanation of it.

Perhaps advocates of the semantic view of scientific theories might contest the connection between inferences and explanation being assumed, most likely on the grounds that, on the semantic view, models are regarded as extra-linguistic entities, while inferences must involve linguistic components. But this technicality will not matter to what follows. I seek to portray selection theory as a set of rules, rules that I think are quite naturally called inference rules. But even if the use of some logical machinery to calculate system dynamics does not count as inference, the deployment and manipulation of formal mathematical structures is surely rule-governed, and that is what matters. I propose simply to regard rule-governed formal manipulation of mathematical structures as a type of inference, if only for ease of presentation in what follows. Besides, the rules with which I will be most concerned are ones that state under what circumstances it is appropriate which models, rather than with the inferences made using the formal machinery.

Selection theory, then, provides a general way of making inferences about the sorts of systems to which it is applied. Accordingly, it must embody a set of *general inferential rules* for making explanatory inferences about the spread of variations through populations. I choose to present selection theory using nothing stronger than the notion of inferential rule because I embrace an approach to philosophical analysis that uses the notions of rule and norm as fundamental explanatory devices. But not everyone will share that broader theoretical perspective, so some discussion of its merits for the project at hand is necessary. Here are a couple of reasons why I think it is beneficial, or at least harmless, to approach selection theory as embodying a set of inference rules.

While selection theory is typically presented as consisting in principles, or models, or laws of nature, rather than rules of inference, thinkers who do not share my philosophical perspective should have little difficulty translating what I write about the rules of selection theory into those terms with which they are more comfortable, or at least more familiar. Laws of nature (principles, models, etc) provide us with rules of inference. Pragmatics aside, would someone really take it that a law of nature does not provide grounds for making inferences about system dynamics? This makes it possible to evaluate a proposed law of nature (or whatever) on the basis of its suitability as a rule of inference. A true law of nature should not induce someone to make incorrect inferences about empirical matters; it should not take them from truths to falsehoods. So my choice to present selection theory as a set of rules should turn out to be harmless from the perspective of those who would prefer more muscular vocabulary. Readers are encouraged to evaluate what I propose as rules of inference for selection theory as laws of nature, or principles, or models, or whatever, whenever it suits them to do so.

Secondly, presenting selection theory as embodying a set of inferential rules rather than natural laws or models allows me to avoid terms that are ambiguous and hotly contested notions in the philosophy of science. Even the rules from selection theory that are most naturally interpreted as expressing laws of nature, such as the inference licenses embodied in the equations of population genetics, will not count as laws of nature according to Earman, Roberts, and Smith (Earman, Roberts, and Smith 2002; Earman and Roberts 1999). These authors are the staunchest contemporary advocates of the view that a law of nature must be exceptionless and universal. In this hardline approach, only the inferential rules of fundamental physics express laws of nature. It will come as no surprise that selection theorists do not and cannot make the inferences they make about the dynamics of natural (though often imaginary) populations by means of the laws of fundamental physics. So presenting selection theory as consisting in inferential rules allows me to present selection theory in a manner that is acceptable to a wider audience, including both those who accept and those who deny that the rules of selection theory should be counted among the laws of nature.

In some ways, I am picking my battles by avoiding ambiguous or controversial alternative language that could be used to present selection theory. But I also mean to expose the value of a normative, rule-based approach to the study of a scientific theory. One can get a lot done, and do a lot of explaining, without using more determinate notion than that of inferential rule to do the work that that notion does here, though I recognize that the proof of this assertion is the pudding.

In the body of the work, I draw close attention to what must be inferred on the basis of what else. My greatest difficulties with rival theoretical presentations of selection theory on the part of both biologists and philosophers have to do with an insufficient attentiveness to the

inferential relationships among the components of the theory, and how these inferential relationships have consequences for the possibility of using selection theory to generate non-circular explanations.

1.4 THE DEFINITENESS OF SELECTION THEORY AS A PRECONDITION FOR EXPLICABILITY

I present selection theory as a definite body, a single, though intricate and complex, inferential tool that can be used to generate explanations of the dynamics of a subset of systems in the natural world, ones whose members are competitors. I present selection theory as having not only a definite function but also a definite content. That content consists in a series of rules for deploying the theory such that it ultimately yields systems of equations that can explain the dynamics of such systems.

If we introduce imaginary or hypothetical populations into this framework, we get accounts of what would happen to such systems. If it is plausible that the hypothetical populations are ancestors of contemporary ones, the theory yields an account of how contemporary populations could have come to instantiate the present day features that they do. It is through these sorts of real deployments of the theory over imaginary systems that demystification is achieved.

The motivation for laying out selection theory as a definite body has to do with how it functions to demystify. For adaptation, say, to be explicable in terms of selection theory, it must be possible to invoke how selection theory could or would be applied to a hypothetical population without actually applying it to a natural system. If we cannot circumscribe selection theory as a specific sort of thing, then our claim that some particular adaptation is explicable

using selection theory will amount to little more than the claim that the adaptation is explicable *somehow*. That sort of assertion does little to demystify. But if instead we can point to a definite set of explanatory techniques that are applicable in definite situations, of which our mysterious adaptation is an instance, the claim that the adaptation is explicable in terms of the theory will have real bite.[2]

1.5 GENERALITY

That selection theory can be deployed over systems that differ in important respects, ways that matter to their dynamics, presents a challenge for anyone who would give the theory a definite content. That selection theory may apply to different sorts of systems that undergo different sorts of dynamics then presents a further line drawing challenge: if the theory can handle a diversity of systems, how do we know which systems are those that undergo selection? Since the theory already encompasses a diversity of systems, how can other, perhaps merely slightly different sorts of systems, be excluded from its scope? On what grounds do we exclude systems that share some features with those considered to undergo selection? If we plump for a particular rule to decide this issue, how do we know whether the rule has too narrow or too broad an application?

I propose to set aside this question for the moment and to return to it in the conclusion. We need not decide the question upfront, and it will be easier to arbitrate it later on, when the presentation of selection theory being pursued here has been completed. It seems harmless to put off the generality issue for a couple of reasons. For one thing, a generalization of selection theory that falls short of full generality need not be considered a total failure, especially if it can be further generalized in the future. Moreover, the approach taken to generalization in this work has

[2] While selection theory should be applicable in each and every instance of what looks like an adaptation, this attitude need not commit us to the view that everything that looks like it resulted from adaptive evolution actually did so.

a broader mandate than just its generality. As I discuss at greater length in section 1.10 below, I use causal vocabulary to generalize the vocabulary found in classical population genetics and by doing so I provide grounds for thinking that selection theory has an explanatory structure. My use of causal vocabulary to present the theory is generalizing, but it also exposes how the theory explains.

It should be noted, too, that my view that it is possible to decide on some definite grounds whether selection theory is applicable to some system is a mainstream view, though few other writers would put the point quite in the fashion I do. At least one strand in the units of selection debate is about the requirements for selection: Dawkins claims that selection requires replicators and calls them "units of selection." Though they sharply disagree with Dawkins on this point, Lewontin and advocates of DST have also invoked a set of requirements for selection (Lewontin 1970, 1978; Griffiths and Gray 1997; Griffiths and Gray 1994; Griffiths and Gray 2001; Griffiths and Gray 2004; Griffiths, Gray, and Oyama 2001; Griffiths and Knight 1998). Brandon (1990) has engaged in the project of saying how critical components of selection theory are interrelated and how they should be deployed, especially "adaptation" and "environment." So, my reasons for doing so aside, that I regard selection theory as a definite body featuring concepts that must be deployed in definite fashion in the generation of explanations of system dynamics is not something that separates me from my interlocutors.

1.6 CIRCULARITIES

One aspect of showing how selection theory is an explanatory theory is showing how the inferences involved are of the explanatory sort. That is a burden taken up later. But something further follows from the claim that explanations are inferences: bad inferences are bad

explanations. Anything that undermines a putative inferential relationship will by extension undermine whatever explanatory power that inferential relationship might be supposed to have. The specific kind of undermining feature with which I will be especially concerned in this section is circularity. Circular inferences do not explain. Hence, an account of selection theory had better not be circular if it is to expose the theory as an explanatory theory.

In the approach taken here, the constraint that selection theory must not embody circularities becomes quite severe, for selection theory is being treated as a complex set of inferentially interrelated rules. The ban on circularity applies not only to individual rules that make explicit, say, how to circumscribe populations or when to weight relative frequency variables with fitness coefficients, but also to the theory as a whole, thereby banning sets of rules that put in place circularities among them. For instance, one cannot make an even implicit appeal to the notion of population in setting down a rule for deploying fitness coefficients if one's rule for circumscribing populations makes appeal to the notion of fitness. Sets of rules that exemplify circularities cannot contribute to an understanding of how selection theory as a whole should be deployed as an explanatory tool.

Circularities also appear in accounts of selection theory in which the explanatory output of the theory is used to make sense of elements internal to it. This particular sort of circularity merits discussion if only because of the large literature that has built up around the "tautology problem." When we define fitness in terms of relative reproduction rate, we undertake this sort of circularity, at least if we mean our definition of fitness to be relevant to the deployment of fitness coefficients in mathematical models. When we define drift as random change in allele frequency, we do the same sort of thing.[3]

[3] I will offer alternative ways to understand fitness coefficients and drift in chapters 6 and 7.

The last sort of circularities I confront are ones that connect statements of the requirements for selection with other elements of the theory, either the explanatory output of the theory, or theoretical terms from the body of the theory. So, for instance, we should not make evolution, or differences in relative reproduction rates or relative fitnesses a requirement for selection. I criticize some candidate statements of the requirements for selection on the grounds that they involve one in circular reasoning of this sort.

1.7 DISTINGUISHING NECESSARY AND SUFFICIENT CONDITIONS

I have promised a philosophical account of selection theory, a theory that makes explicable otherwise mysterious phenomena, chiefly adaptation, genetic polymorphism, and altruism. That selection theory must make mysterious phenomena explicable means the theory must be structured in an appropriate fashion to do so. In other words, that selection theory demystifies adaptation implies further facts about the theory. These features are the *necessary conditions* of the theory. Anything that does not show how selection theory has these features is simply not recognizable as an explication of selection theory. Conversely, anything putatively sufficient to meet these obligations deserves consideration as a candidate explication of selection theory.[4]

[4] I recognize that this commits me to regarding a strange amalgam of what I call selection theory and another unrelated theory such as quantum mechanics as an instance of selection theory. Such things should not be counted as instances of selection theory, but acknowledging them as such here just makes my life more difficult, for it increases the number of theories I am forced to consider as rivals, so it should not bother us for the present. I come back to discussing this sort of amalgam in the conclusion (chapter 8).

I summarize and review the necessary conditions of selection theory. For something to count as selection theory, it must be a fit tool for making our trio of mysterious phenomena explicable. That necessary condition begets several more requirements in turn. First, selection theory must have an explanatory structure and (consequently) an inferential or at least rule-governed one. Second, it must not violate strictures on what is a good sort of inference. Particularly, it had better not be somehow circular. Last, selection theory must be isolatable as a definite body in order to fulfill its function of demystification.

The remainder of this work consists in an account of selection theory that shows how the theory has the features just reviewed. What I offer as at least *a version* of selection theory meets the necessary conditions that an account of selection theory must meet. Indeed, I justify my rules for how selection theory should be understood in terms of how they allow one to accomplish the *point* of the theory, the demystification of features of the organic world, especially adaptation and adaptive evolutionary change. My suggestion for how selection theory is structured fulfills the necessary conditions for an account of selection theory, it has the various features that any such account must have: it is inferential, explanatory, non-circular, and can be isolated as a definite body.

Justifying a proposed understanding of a theory, or indeed an understanding of any fragment of language, on the basis its point involves an appeal to pragmatics, an appeal to the interests of individuals who would deploy it. It is only those of us who are interested in understanding what could explain adaptations and other curious phenomena who can get something out of an understanding of the rules that follow. But the rules of reasoning I offer below apply only to those interested in making adaptation explicable.

I am not the only person interested in the functioning of selection theory who argues in this way. Another writer who does so is John Damuth (1985). Concerned to argue that clades are unsuitable as populations over which to deploy selection theory, he repeatedly argues against so treating them on the grounds that selection theory is not "designed" to study such populations, that it would not "effective way to study" clades and that doing so would result in an "undesirable alternation in the explanatory scope of the theory" (Damuth 1985, 1134`; 1136).

Clades are spread out over great distances and include organisms that do not interact with each other or with similar environments. These are certainly not the sort of populations over which selection theory is typically deployed. However, Damuth is clear that there is no logical bar to a version of selection theory in which clades function as populations, it's just that such a theory would not yield the sort of explanations that we are after when applying selection theory:

> The overall effect of treating changes in non-localized entities [such as many clades] as results of a pure, simple selection process operating on each entity as whole would be to drastically alter the sorts of questions about causation that could be addressed by this modified theory. In conventional simple selection, fitness values and observed or predicted changes represent the effects of the *relative quality* under given circumstances, of different phenotypes in the population. In this new, expanded usage of simple selection, this would no longer be the case. … But, it is the conventional theory that provides the explanations of interest to us. We have no use for a theory of pure selection that explains or predicts that a population of type A, experiencing particular conditions in Africa, will shrink, while a population of type B, experiencing different conditions in North America, will grow—and that thus the proportions of the two types will change. Rather, we want a selection theory that deals with the relative abilities of the two types under particular circumstances, and we want it to tell us how these relative abilities cause populations in each place to change in response to local conditions (i.e. to adapt). The more complex, global picture involving change in the non-localized entity is a job for a different kind of model (Damuth 1985, 1134; italics in original)

Note how Damuth repeatedly makes reference to what we expect from our models, what their jobs are, what they tell us about what interests us. Below, I repeatedly make reference to the point of selection theory, to what it is supposed to tell us, to justify my decisions about how to understand the theory. These justifications only work on individuals who are interested in acquiring what selection theory provides for us, the demystification of adaptation, genetic polymorphism, and altruism.

It should be stressed, too, that what is on offer below is nothing more than one suggestion for how the point of selection theory might be fulfilled. What I offer is not logically compelled by the necessary conditions that an account of selection theory must fulfill. An account of selection theory must make it clear *how* one proposes to meet the explanatory obligations of selection theory, how to fulfill its point. But to show one way of doing so is not to show the only way to do so. This section thus forms a bridge: the features of selection theory discussed earlier are the necessary conditions that any account of selection theory must meet, what comes next are a set of sufficient conditions for meeting these.

1.8 ENTRANCE RULES, EXIT RULES, AND THE BODY OF SELECTION THEORY

In this section I will make a little more precise my ideas about how selection theory functions to demystify. The discussion that follows is placed firmly in the sufficient conditions segment of the dissertation. It might well be possible to fulfill the requirements that a presentation of selection theory must fulfill in some alternative fashion. What follows is *one way* to present selection theory as an explanatory theory. The strategy I adopt is neither that of the semantic nor the syntactic view of scientific theories. For the sake of having a label, I will call my approach inferentialist.

As presented here, the theory consists in three sorts of inferential rules. The first sort of rule I call an *entrance rule*. An entrance rule is a statement of what sorts of features some system must have for it to be correct to deploy selection theory over that system.[5] Equivalently, an entrance rule states the circumstances of application of selection theory. In chapter 2, I consider a number of statements by other writers concerning the requirements for selection. I evaluate, and ultimately reject, these as candidate entrance rules for selection theory. While the authors of the statements of the requirements for selection I consider would probably not have called their statements "entrance rules," it remains a widespread intuition among philosophers of biology and biologists that it should be possible to say what sorts of features some system must have for it to make sense to treat its dynamics using selection theory. I offer my own stance on the entrance rule for selection theory in chapter 3.

Skipping ahead, the sorts of rules that come last in selection theory are the *exit rules* of the theory. The exit rules of the theory are systems of equations or mathematical models, ones that are constructed in the manner of classical population genetics. I speak of exit *rules* rather than a single exit *rule* because there are many mathematical structures that are used to model the dynamics of populations under selection.

The last set of rules, the ones that make up the body of selection theory, consists in rules for connecting a system that meets the entrance rule for the theory to some particular exit rule, a mathematical model that governs the dynamics of the system. Selection theory has a simple entrance rule, one that sorts systems into two piles, those over which it makes sense to deploy the theory and those it does not. Somehow, the systems over which the theory is applicable must be further sorted so that it is possible to say which ones are associated with which mathematical

[5] By using the terms entrance rule and exit rule, I am deliberately trying to invoke the Gentzen-style approach to the introduction and elimination of logical vocabulary. The idea behind my presentation of selection theory is much the same, though I am talking about the deployment of a theory, not a term, a concept, or a logical operator.

structures among the exit rules of the theory. There are rules for making these sorting decisions, indeed a quite intricate set. Those rules constitute the body of the theory; they are what come between the entrance and the exit rules and link the two.

Though it should in general be possible to state which mathematical structures are appropriate for which sorts of systems, I do not attempt to do this for the entirety of population genetics. Instead, I focus my attention of sub-class of mathematical models, Wright-Fisher discrete generation models. I have chosen to deal with discrete generation models for a couple of reasons. Discrete generation models are mathematically the least sophisticated of population genetics models, making them the easiest to understand. But this lack of mathematical sophistication brings with it the possibility of modeling a great variety of causal scenarios, far more than can be handled by diffusion theory, Cannings models, Moran models, age-structured models, or a number of other alternatives.[6]

[6] Indeed, philosophers interested in evolutionary theory have often failed to get a grip on how the sophisticated models of selection theory can accommodate the existence of contextual causal influences over population dynamics (e.g., Glymour 2006)

As part of my discussion of the body of selection theory, I first offer, in chapter 4, an account of how to group the entities over which the theory may be applied into populations. In chapter 5, I offer some definitions for use in an algorithm for making judgments about the appropriate deterministic Wright-Fisher model for an arbitrary population. In chapter 6, I present the algorithm for generating a deterministic Wright-Fisher mathematical model appropriate for some population based on causal facts about it. In chapter 7, I briefly show how to deploy stochastic versions of Wright-Fisher models by taking into account random influences over dynamics, ones that the deterministic versions generated by the algorithm of the sixth chapter leave out. Those chapters together comprise an account of how to connect systems over which selection theory can be applied to mathematical models of their dynamics, at least for a considerable subset of population genetics models.

1.9 SELECTION THEORY AS A CAUSAL THEORY: EXPLANATION

I have claimed that anything that counts as selection theory must have an explanatory structure, for it must demystify adaptation. One way a theory can earn its credentials as an explanatory theory is for it to be a *causal theory*. While equations and principles that expose correlations might make reliable inference rules, ones that establish causal relations are explanatory.

Merely demonstrating that selection theory can be used to draw conclusions about the dynamics of populations is not enough to show that the theory is explanatory because some phenomena may be inferred without being explained. The standard contemporary example of an inferential relationship that is non-explanatory is the flagpole/shadow case: the length of the shadow of a flagpole, along with some additional information, can be used to infer the height of the flagpole itself, but the length of the shadow hardly explains the height of the pole. Most would say the reverse is true.

As presented here, selection theory culminates in exit rules that consist in causally interpretable equations. In these equations, variables picking out individuals distinguished by their genetic variations are functioning as causes and for this reason the equations should be regarded as explanatory. That causation is at least sufficient for explanation is a stance endorsed by several recent writers on the topic (Salmon 1998; Woodward 2003). We will be concerned to demonstrate how to infer the deployment of one or another system of causally interpretable equations for some system.

One way to secure a causal interpretation for systems of equations is to infer them from causal graphs, graphs that are themselves produced on the basis of statistical data about a system coupled with a variety of assumptions (Glymour, Scheines, and Spirtes 1993; Pearl 2000; Woodward 2003). I do not follow this approach. Instead, I make critical use of causal facts in an algorithm that yields as its output a system of equations that can be used to calculate system dynamics. The use of causal information in the algorithm is what is supposed to show that the resulting equations are causally interpretable. The algorithm does not take advantage of nothing but causal information; I also make use of mathematical, and more generally logical vocabulary,

along with statistical vocabulary. However, it is the critical contribution made by causal vocabulary that underwrites my claim that the algorithm yields systems of equations that explain.

While I claim that the equations of population genetics explain because they are causally interpretable, I do not claim that causation is necessary for explanation. For one thing, there might be moral or mathematical explanations, but also, by defining terms one can explain them, though such explanations are clearly not causal in character. So, showing how selection theory functions as a causal theory is not necessarily the only fashion one might demonstrate that it functions as an explanatory theory; it is merely the line pursued here.

A further issue concerning the relationship between causation and explanation merits consideration. The claim that causal relationships are explanatory might invite a reductio: there are seemingly innumerably many causes of a particular event, many of which do not strike us as explanatory, so causation cannot be sufficient for explanation. For instance, the Hindenburg would not have burnt up had there been no oxygen in the atmosphere, so in some sense at least, the presence of oxygen is a cause of that disaster.

If we refine our interest a little and ask why the Hindenburg burned to the ground rather than landing safely, then we can rule out oxygen as a cause of its burning, since the Hindenburg would not have landed safely in a world without oxygen; indeed it would never have been built or piloted by humans who rely on oxygen to breathe. This sort of contrastive thinking about why one thing happens rather than another can help focus our explanatory interests so as to exclude non-explanatory causes. To turn to the case at hand, when we deploy selection theory, we are interested in explaining why one rival type achieved the relative frequency that it did within some population rather than achieving a different frequency; many causes, such as the non-destruction of the Earth by a meteor, cannot help us explain why the population arrived at one

relative frequency rather than another. Furthermore, some causes of population dynamics are *null causes*, ones that, were they represented in causally interpretable equations, could be dropped from the equations without this having any impact on calculated quantities. Null causes need not be regarded as explanatory either (for a fuller discussion of null causes, see Gildenhuys (forthcoming)). So, officially, my stance that causes explain is really more specific: non-null causes that allow us to infer why one member of a contrast class of eventualities of interest occurred rather than another member of the contrast class are explanatory. In classical population genetics, the contrast class is easy to pick out, since the equations are used to infer relative frequency terms, the contrast class includes all the possible relative frequencies, including those that eventuated and those that did not

My reasons for taking causation to be sufficient for explanation are those of Woodward; I really have nothing to add to Woodward's argument that causally interpretable equations are explanatory (Woodward 2003, ch. 5). In the remainder of this section, I merely present a summary of the Woodward account of causation and discuss it; I do not argue for it.

Woodward's definition of causation is rather involved, though as he notes himself, it is less complicated than it looks. Here's the definition:

> (**M**) A necessary and sufficient condition for X to be a (type-level) direct cause of Y with respect to a variable set V is that there be a possible intervention on X that will change Y or the probability distribution for Y when one holds fixed at some value all other variables *Zi* in V. A necessary and sufficient condition for X to be a (type-level) *contributing cause* of Y with respect to variable set V is that (i) there be a directed path from X to Y such that each ink in this path is a direct causal relationship; that is, a set of variables Z1 … Zn such that X is a direct cause of Z1, which in turn is a direct cause of Z2, which is a direct cause of … Zn, which is a direct cause of Y, and that (ii) there be some intervention on X that will change Y when all other variables in V that are not on this path are fixed at some value. If there is only one path P from X to Y or if the only alternative path from X to Y besides P contains no

> intermediate variables (i.e., is direct), then X is a contributing cause of Y as long as there is some intervention on X that will change the value of Y, for some values of the other variables in V. (Woodward 2003, 59)

It should be noted that Woodward's definition makes reference to interventions that are themselves defined in causal terms. Here is Woodward's definition of an intervening variable (*I*) for *X* (a possible cause) with respect to *Y* (a possible effect of *X*):

> I1. *I* causes *X*.
> I2. *I* acts as a switch for all the other variables that cause *X*. That is, certain values of *I* are such that when *I* attains those values, *X* ceases to depend on the values of other variables that cause *X* and instead depends on the value taken by *I*.
> I3. Any directed path from *I* to *Y* goes through *X*. That is, I does not directly cause *Y* and is not a cause of any causes of *Y* that are distinct from *X* except, of course, for those causes of *Y*, if any, that are built into the *I*-*X*-*Y* connection itself; that is, except for (a) any causes of *Y* that are effects of *X* (i.e. variables that are causally between *X* and *Y*) and (b) any causes of *Y* that are between *I* and *X* and have no effect on *Y* independently of *X*.
> I4. *I* is (statistically) independent of any variable *Z* that causes *Y* and that is on a directed path that does not go through *X*.

Interventions, then, get defined this way:

> (IN) *I*'s assuming some value $I = z_i$, is an intervention on *X* with respect to *Y* if and only if *I* is an intervention variable for *X* with respect to *Y* and $I = z_i$ is an actual cause of the value taken by *X*. (Woodward 2003, 98)

Woodward's definition makes explicit mention of node-and-directed-edge graphs popular in contemporary thinking about causation (Glymour, Scheines, and Spirtes 1993; Pearl 2000). Later on, I use such graphs to offer causal representations of what is going on in systems to which selection theory models are applied. With such graphs at hand, it is easy to check what causes what. More importantly, once they are coupled with functions to represent the extent and

character of the dependencies among the variables that they portray, such graphs can be used to secure the causal interpretation of the mathematical equations that can be extracted from them.

1.10 CAUSATION AS A PRIMITIVE

Though I regard Woodward's construal of causation as valuable, I do not consider it an explanatory account of that notion. For one thing, the fact that the definition makes appeal to causal notions means it cannot supply a general understanding of what we are to treat as causes. We must already understand and know how to deploy the notion of causation to use Woodward's definition to get a grip on whether or not two variables are causally related. We cannot in general infer causal relations using Woodward's definition, so we cannot in general explain what causal relationships are using it either.

Another similar reason to refuse to regard Woodward's definition as explanatory is that we cannot in general use causal graphs to pin down causal facts, so we cannot in general explain what it means to be a cause by appeal to them. We generate causal graphs in two ways, but neither way provides us with a general way of understanding what causes are. We use causal facts as inputs for causal graphs, as we do when we teach people how to use graphs to portray causal relationships. In these cases, the graphs cannot be used to explain the causal relations; rather, the reverse is going on. We can also generate causal graphs using algorithms for causal inference of the sort offered by Pearl, Scheines, Glymour and Spirtes and their students. These take statistical facts as their input, but unique causally interpretable graph structures are not always implied by statistical data. Not only do the algorithms sometimes take advantage of assumptions, such as faithfulness, which need not hold of some causal system, but they also sometimes yield patterns, graphs that do not pick out a specific causal structure as what the evidence implies.

Generally, it is not the case that there exist causal relationships among a set of variables whenever we can arrive at graphs to represent those relationships. The causal relationships exist in the system even when we cannot use any particular algorithm to decide what they are. So we cannot make sense of what it means to be a cause by appeal to our abilities to draw causal graphs, just as we cannot use facts about the behavior of scales that measure weight to explain what it means to say of objects that they have mass. Even in cases where we cannot compose a graph for some data set that shows the causal connection between two variables, a causal relationship between them may nonetheless exist, just as an object may have mass even when we cannot weigh it on any particular scale.

In light of this, and the circularity of Woodward's definition, I propose to treat causation as a primitive theoretical term in the same way that "mass" or "force" are theoretical terms in classical physics. "Cause" for me will function as an unexplained explainer. We can sometimes put ourselves in a position to affirm causal relationships on the basis of evidence, and that is what algorithms for causal inference allow us to do, but we do not have a hard and fast rule, versed in some more basic vocabulary, for deciding what counts as a cause, and hence we cannot use a rule, an algorithm or a definition, to explain the circumstances of application of "cause."

The situation matches that of fundamental physics where the deployment of values for theoretical terms can be established in specific cases on the basis of evidence. Both claims about causal relationships and claims about unobservables have implications for system dynamics, but the meanings of some of their crucial bits of vocabulary cannot be specified in antecedently available vocabulary (Suppe 1974). So the fact that a scientific theory should harbor theoretical terms that cannot be defined or otherwise discharged in favor of a more basic sort of vocabulary, such as the language of observation, should not surprise us. Though the undischarged vocabulary is different in each case, being versed in undischargeable vocabulary is something that fundamental physical theories and causal theories, such as selection theory, share.

One might worry that my stance is too strong here, and that one ought to regard the Woodward definition of causation as providing an explanatory account of causation. Equally, one might worry that the algorithms used to generate causal graphs from statistical data really should be regarded as providing a robust account of how the notion of causation is deployed. I have argued against both views, but there is little point in pursuing a debate with those who would take either of these stronger views about causation. Those who would take up the stance that "cause" can be defined in an explanatory fashion, and those who believe that algorithms can

be used to provide the circumstances of application of causal relationships, will find nothing amiss with my use of causation as a key explainer. Instead, they will simply take the view that what are here treated as primitives can be explained in even more basic vocabulary.

Although treating causation as a primitive in this way gives up a lot of ground, I see no alternative to doing so. Causation is, anyhow, a notion with which any well-educated competent English speaker is familiar. Indeed, Woodward repeatedly takes advantage of this fact to argue against competing accounts of causation and explanation. For instance, because we are already willing to grant that there are cases of causation by omission, but which do not involve energy transfer, we should count it against an account of causation as energy transfer that it cannot accommodate such causal relations (Woodward 2003, 91).

Officially, then, selection theory can be used to explain population dynamics because it yields causally interpretable equations that provide a means of inferring those dynamics. In the account of selection theory that follows, equations for population dynamics will be assigned to specific systems on the basis of an understanding of the causes operative in them. But the explanation I offer of how to model population dynamics hits rock bottom at the level of causation: I do not try to say when we should regard what systems as exhibiting which sorts of causal relationships. It should be noted, too, that I present the Woodward construal of causation as the preferred understanding of that notion, but little hinges on this. Any alternative account that makes the equations of population genetics come out causal will do just as well.

1.11 SELECTION THEORY AS A CAUSAL THEORY: HISTORY

It is a lead claim of this work that the theoretical vocabulary from selection theory can be defined; specifically, theoretical terms in selection theory can each be defined by making use of

the notion of *causation*. Explications of the meanings of critical terms in selection theory, including the variables that refer to different types of individual, can be constructed out of a regimented version of the causal vocabulary of ordinary English, indeed, the sort of English available to biologists and other educated persons of both the mid-Nineteenth Century and today.

Selection theory is a causal theory in a double sense, then. While the equations of population genetics that form the exit rules of selection theory are susceptible to causal interpretation insofar as they feature causes of the right and effects on the left, its theoretical terms are susceptible to causal explication, too. They are either susceptible to causal definition, as is the case with population, or their deployment can be specified as part of the algorithm that takes causal facts as inputs and yields equations featuring the theoretical terms as outputs.

That selection theory can be understood by making critical use of the notion of causation as an explainer fits well with the history of its inception. Understanding the variables of selection theory as getting their meanings by way of the notion of causation helps to understand how selection theory can at once contain no novel undefinable vocabulary referring to novel unobservable entities, forces, or the like, and yet also be a theory whose articulation by Darwin was a tremendous scientific achievement, one that required great insight. As a concluding section to this preliminary chapter, I will argue that the approach pursued here jibes well with the history of selection theory as first put forward by Darwin. Specifically, I will be concerned to show that the causal explication of selection theory pursued here fits more naturally with its history than an account of selection theory according to which it gets it explanatory power from positing novel unobservable phenomena.

Selection theory is not obvious. An understanding of selection theory involves recognizing how an array of features of (paradigmatically) organic populations can lead to

evolutionary change (or other sorts of dynamics) within them. Yet despite the fact that selection theory is not obvious, Darwin's *Origin of Species* contains none of the sort of discussion one would expect from a founding work in a scientific discipline that brings to the attention of the scientific community the existence of previously unobservable or unknown phenomena with novel explanatory power. Selection theory is unlike fundamental physics and chemistry in this respect; something that is clear from the character of its founding text. Darwin did not posit novel unobservables as explainers but instead drew out previously unrecognized implications of causal relationships.

Prior to the publication of *The Origin*, Darwin was at work applying his theory of natural selection in an effort to form a broad inductive base for his assertion that evolution occurs and is explicable in terms of the process of natural selection. Though Darwin was dismayed by his reception of Wallace's letter in June, 1858, in which Wallace proposed a theory similar to Darwin's, he nonetheless wrote, in a letter to Lyell, that his work applying the theory would not go to waste: "So all my originality, whatever it may amount to, will be smashed. Though my Book, if it will ever have any value, will not be deteriorated; as all the labor consists in the application of the theory" (letter to Lyell, 18 June 1858).

Wallace's letter prompted Darwin to postpone his project of applying his theory to natural populations in favor of publishing the *Origin*, a much shorter work than the one in which he had initially planned to introduce his theory, and one notably lacking in just the sort of evidence of the operation of natural selection in nature that Darwin had been busy gathering. As James Lennox has argued about Darwin's argument strategy in the *Origin*, "whatever else Darwin was doing in the first four chapters, he was not establishing that natural selection had operated as he claimed in nature" (Lennox 1991, 226). It is in the first four chapters that Darwin

establishes the existence of natural selection, and there Darwin does not even offer the right sort of evidence in the *Origin* to mount an explanationist defense of the existence of something new with novel explanatory power.

This is not to say that selection theory cannot be used to generate novel predictions and explanations, nor is it to say that an explanationist defense of the veracity of the theory of natural selection is impossible. Rather the successful deployment of selection theory to explain natural phenomena was not what was supposed to convince Darwin's readers of the existence of natural selection. Instead of demonstrating the explanatory power of his theory by deploying it over natural systems, the sort of thing that would function to justify novel theoretical posits therein, Darwin argues for the existence of natural selection by way of a Herschelian generalization of the artificial selection, a principle already understood to operate upon organisms under domestication.

As I have discussed more fully elsewhere (Gildenhuys 2003), Darwin's argument for the existence of natural selection is a step-by-step generalization that extends the principle of selection from domestic populations to natural populations. Having cemented in chapter 1 how artificial selection explains the bulk of evolutionary change in domestic organisms by way of the accumulation of variations in one or more directions, Darwin goes on to show how other practices institute selection regimes, too, each one departing a little bit further from the base case of artificial selection. Darwin discusses roguing, the removal from seedbeds of deviant forms, the import of choice animals from foreign countries for breeding, and the "unconscious selection" of owners who breed only their best animals without trying to create forms with specific traits. He then considers domestic animals kept by "savages" who are too unsophisticated to control which of their animals breed. Animals in their stock nonetheless differentially reproduce, despite the

inattention of their masters, because of their variant constitutions and structures, a process leading to evolution. It is a short step from this last scenario to evolution by selection in nature; indeed, Darwin regards the domestic animals kept by inattentive savages as evolving by natural selection despite the fact that they live in under domestication (Darwin 1988[1859], 38).

Darwin's generalization of the process of artificial selection is a matter of exposing how the *same* causal process that goes on under domestication can go own in nature too. Selection theory does not explain by positing the existence of previously unknown and unobservable forces or entities, ones whose properties, relations, and activities can be shown to result in the sorts of things that selection theory is designed to explain. Rather, selection theory explains by exposing how a process already well-understood in one domain (domestication) can also occur in another domain (nature).

It should be noted that while Darwin does not establish the existence of natural selection by way of an explanationist defense of the value of it as a theoretical posit, Darwin does not in general shun explanationist modes of reasoning. As Philip Kitcher (1985) has argued, Darwin's argument to the effect that evolution by natural selection does a better job than creationism of explaining biogeography, comparative anatomy, and embryology, is very much explanationist in character:

> Darwin's approach [in the later chapters of the *Origin*] is to marshal an impressive array of puzzling cases of geographical distribution, affinity of organisms, adaptation, and so forth, aiming to convince his reader that there are numerous questions to which answers hitting his schemata would bring welcome relief. (Kitcher 1985, 150).

Though he *does* use an explanationist strategy to demonstrate that evolution is a better account than creationism of a diverse collection of facts, Darwin does not use an explanationist approach early on in the *Origin* to justify the *existence* of natural selection.

For the sake of a contrast, consider how Darwin's extension of previously understood processes and principles to novel domains contrasts with Mendel's argumentative strategy in his essay "Experiments in Plant Hybridization." Mendel argues for the existence of "differentiating elements" of egg and pollen cells that are tied to variable characters on the grounds that the existence and activities of these elements can explain patterns of hybridization in his experimental plants. Mendel offers his theory as a hypothesis, indeed a tentative one requiring further testing:

> The attribution attempted here of the essential difference in the development of hybrids to a *permanent* or *temporary union* of the differing cell elements can, of course, only claim the value of an hypothesis for which the lack of definite data offers a wide scope. Some justification of the opinion expressed lies in the evidence afforded by *Pisum* that the behavior of each pair of differentiating characters in hybrid union is independent of the other differences between the two original plants, and, further, that the hybrid produces just as many kinds of egg and pollen cells as there are possible constant combination forms. (Mendel 1901[1865])

Unlike readers of Mendel's work, who must acquire a novel piece of vocabulary, "differentiating cell element," with novel implications for the distributions of traits in hybrid crosses, those who learn selection theory do not need to learn new words for new things, ones with novel influences, effects, and activities, and ones of whose existence they were previously unaware. This is true even for "natural selection."

While "natural selection" is certainly a novel idea, its existence is established by Darwin by means of a Herschelian analogy, not by its ability to explain. Natural selection is best conceived as a process (Hull 1988), and while this process can explain the evolution of species, the process itself is explained in other terms. For Darwin these are variation, inheritance, and the struggle for existence, conditions that were antecedently understood to obtain in nature. Where

such conditions obtain, a process of selection is instantiated, itself a sort of process also previously understood to go under domestication (making that process a *vera causa*).

Indeed, it is not at all clear what one might choose to play the role of a theoretical term in selection theory were one to reconstruct selection theory as a theory involving novel entities picked out by antecedently unused terminology. As will become clear from the discussion later on, the most inferentially powerful idea in Darwin's theory is that of the struggle for existence in natural populations, but Lyell had already discussed that phenomenon at length well before Darwin's *Origin* (Lyell 1969[1830], 131-140), and the struggle for existence hardly qualifies as the sort of thing we believe in because of what it can explain. Darwin saw farther than Lyell into the implications of the struggle for existence, but he was not the first to register the existence of the phenomenon in the natural world.

"Fitness" might make a more natural choice as a theoretical term, as its interpretation has been subject to intense debate throughout the history of evolutionary biology. But, as a noun, "fitness" surfaces only once in *The Origin*, and then long after the work presenting the theory of natural selection has been completed. Darwin does formulate claims about different varieties and types within populations as being fitted to their environments more or less well, but these claims are always closely connected to a definite example, one in which observable variations among organisms form the inferential basis for the comparisons. What's more, talk of organisms being more or less fit appears only once prior to the argument establishing the existence of natural selection in a discussion of domesticated sheep (Darwin 1988[1859], 30). And of course, "fitting" better or worse into a broader environmental context is hardly the sort of thing that is unobservable. So the *Origin* gives us not reason to think that "fitness" forms an essential component of the theory of natural selection by picking out a hidden, previously unknown feature of organisms that contributes to the generation of explanations that cannot be conducted in other terms.

Instead of conceiving of selection theory as getting its explanatory "oomph" from positing the existence of novel unobservable theoretical entities, I explore an alternative view. In the account proffered below, selection theory is a theory that brings to light previously unexplored causal relationships, and draws implications from these. The explanatory power of the theory lies in its elucidation of causal relationships among macroscopic phenomena, causal relationships whose impact on the dynamics of populations had gone unrecognized prior to Darwin. Such causal relationships are that to which we must appeal in order to delineate the sorts of entities that play the various causal/inferential roles they do in the theory.

Consider briefly, as an example of a term that requires causal definition, the notion of "population" in selection theory. Darwin did not understand his theory as applying to species. Darwin famously rejected the possibility of distinguishing in some principled fashion between individual variations, varieties, and species (Darwin 1988[1859], 52). Darwin saw a continuum of differences here. Besides, if the process of natural selection is supposed to show how species emerge, it can hardly be a requirement that members of a population be members of the same species for the theory to be applicable to them unless "species" is explicitly taken to be a vague term. What's more, as Damuth (1985) has pointed out, populations circumscribed on taxonomic grounds do not necessarily make for good populations over which to deploy selection theory. So Darwin was right not to delineate the populations that evolve by natural selection using pre-existing ways of carving up the biological world.

Furthermore, selection can go on between members of different species. Consider how, to get his concept of natural selection across, Darwin invites his reader to understand the spread of variant forms within a local population analogously to the parallel process that occurs when immigrants arrive at a territory that has recently undergone a physical change, such as a change in climate:

> If the country were open on its borders, new forms would certainly immigrate, and this would also seriously disturb the relations of some of the former inhabitants. Let it be remembered how powerful the influence of a single introduced tree or mammal has been shown to be. But in the case of an island, or of a country partly surrounded by barriers, into which new and better adapted forms could not freely enter, we should then have places in the economy of nature which would assuredly be better filled up, if some of the original inhabitants were in some manner modified; for had the area been open to immigration, these same places would have been seized on by intruders. (Darwin 1988[1859], 81-82)

As the quotation shows, a selection process can go between rival contenders for the same place in an ecosystem, no matter what the origin of the rival variants. Darwin is quite explicit that members of different species can struggle for existence with one another (Darwin 1988[1859], 63).

The upshot of Darwin's refusal to circumscribe populations suitable for his theory in terms of species membership means that his theory applies to populations delineated in some other way. For Darwin, population members must be engaged in the struggle for existence, something that means they must be engaged in a specific sort of causal relationship with one another. While Darwin does not define this sort of causal relationship, he does discuss it at length in the third chapter of the *Origin*, analyzing what produces the struggle for existence, offering characteristic instances, and distinguishing cases in which two organisms can correctly be said to be struggling for existence in the sense important to Darwin from cases in which they cannot (Darwin 1988[1859], 62-3).[7] I discuss a regimented version of how to circumscribe populations in chapter 3 that takes the form of a definition of "population" in selection theory, a definition that makes critical use of the concept of causation.

Causal relationships are not themselves observable or definable in non-circular terms. So, to use the last example, using the struggle for existence to parse organisms into populations requires seeing the value of doing so, seeing the consequences that the struggle for existence has among entities that vary. Presenting selection theory as a theory whose critical explanatory vocabulary should be understood in causal terms, then, helps explain the history of how the existence of natural selection was surmised.

[7] I note that the sense in which a lone cactus at the edge of the desert struggles is *not* the sense in which creatures struggle for existence.

The remainder of the dissertation deploys cause-talk to make sense not only of how populations are delineated, but also to state an entrance rule for selection theory, to define other variables, and to describe how the various mathematical models of population genetics attach to real world systems. The completed project is a set of rules for applying the theory that allow one, in principle, to understand the dynamics of systems over which the theory may be deployed. An understanding of these rules allows one to understand how adaptation, genetic polymorphism, and altruism are explicable, for it shows how it may be inferred by means of an explanatory theory for systems whose features are described in causal terms.

2.0 THE REQUIREMENTS FOR SELECTION

The explanatory apparatus of selection theory is deployable only upon some systems. This makes selection theory no different from other scientific theories, with the possible exception of fundamental theories of physics. Because selection theory can be deployed only over some sorts of things, it may be possible to pick out those systems explicitly. Indeed, if selection theory is to make adaptation and other mysterious phenomena explicable, it had better be possible to invoke its applicability to some system without actually applying it to that system. So, those who would use selection theory to demystify adaptation, polymorphism, and altruism must have some understanding of the sorts of systems over which the theory may be deployed. Accordingly, we need, or at least should value, some rule to distinguish those systems over which selection theory may be deployed from those over which it may not.

I call the rule that distinguishes between systems over which selection theory may be deployed and systems over which it may not be deployed the *entrance rule* for the theory, since meeting the entrance rule triggers the deployment of the rest of the rules over the system that constitute selection theory. Stating my proposal for the entrance rule for selection theory is the order of business of the next chapter of this work. In this chapter, I will first consider some accounts of the requirements for selection and show that they do not function effectively as entrance rules for selection theory.

While no other author has characterized selection theory as having an entrance rule, plenty of other authors have sought to state the *requirements for selection*. It is not always clear whether other writers' stances on the requirements for selection amount to stances on what I would call the entrance rule for selection theory. Still, for each of the authors I consider below, I make a case that what they have written about the requirements for selection are reasonably interpreted as statements of an entrance rule for selection theory. But whether or not the authors of statements of the requirements for selection intend their statements as what I call entrance rules is really a secondary issue. My main purpose in this chapter will be to show that prominent statements of the requirements for selection currently on offer in the literature fail to function adequately as entrance rules for selection theory. Achieving this purpose will motivate my own proposal.

That the requirements for selection, or the entrance rule for selection theory, should be the subject of philosophical investigation may strike the reader as initially strange. Do not population genetics textbooks specify what systems fit their models? In one sense they do: population genetic models are deployed in the face of genetic variation among biological organisms. However, the suspicion at work here, and one shared by almost everyone who has considered the topic of the requirements for selection, is that genetic variation is not a requirement for selection. An entrance rule for selection theory should be pick out a broader range of systems that just those exhibiting genetic variation as ones over which it is appropriate to deploy the theory.

The attitude that more than just genes undergo selection is shared by a wide range of writers. Griffiths and Gray, advocates of expanded inheritance in the framework of developmental systems theory, suggest that such things as chromatin marking schemes and

habitat imprints should count as inheritance mechanisms on a par with DNA. They even claim that aphids that vary in the endosymbiotic bacteria they harbor may be subject to selection on variant lineages of the bacteria (2001, 198). Dawkins, who shares few of Griffiths and Gray's views about selection, defines the notion of replicator explicitly so as not to "prejudge the empirical issues" concerning whether an entity is of the right sort to undergo natural selection (quoted in Hull 1980, 317; see Dawkins 2004 for his attitude to putative non-genetic inheritance mechanisms). Once the possibility of selection without specifically genetic variation is acknowledged, stating the requirements for selection in a manner that does not let in too much or too little becomes a real challenge and a recognizably philosophical one.

In the remainder of this chapter I lay out some general constraints on what features an entrance rule for selection theory must have. Then, I argue against what I call *resemblance selectionist* accounts of selection in this chapter, ones developed out of Lewontin's (1970) statement of the requirements for selection. In the next chapter, I develop my own *type selectionist* stance on the entrance rule for selection theory alongside a consideration of Dawkins' replicator selectionist approach. If we make the modifications I propose to Dawkins' view that selection requires replicators, we end up with my own view that selection requires *competitors*.

2.1 CONSTRAINTS ON AN ENTRANCE RULE FOR SELECTION THEORY

Lewontin is the author of the most widely referenced statement the requirements for selection, a set of three principles that together putatively embody the "principle of natural selection" (Lewontin 1970). Though Lewontin's three principles are often modified a little by subsequent writers, Lewontin's text is also treated as authoritative; he is even sometimes said to have provided an argument that his principles set out the requirements for selection (e.g., Godfrey Smith 2000, 2). But Lewontin provides no such thing; rather he simply declares that his three principles embody Darwin's theory (Lewontin 1970, 1). In this section, I offer an account of what sorts of features a set of requirements for selection must have if they are to function as an entrance rule for selection theory. This will provide a basis for arguing over whether some putative set of requirements for selection can actually function adequately as such. By clarifying what we want from an entrance rule for selection theory, we can set out criteria by which to evaluate different proposals.

There are many ways in which a statement of the requirements for selection could fail to pass muster as an entrance rule for selection theory. The constraints on an entrance rule for selection theory include at least these:

- An entrance rule should not be too restrictive. An entrance rule for selection theory is too strong if it restricts the application of the theory to only a subset of those systems over which the rest of the rules of the theory could be exercised so as to generate the sorts of conclusions that we use selection theory to draw. An entrance rule for selection theory that requires the existence of genetic variation is too restrictive.

- An entrance rule for selection should make the theory deployable only over those systems that actually behave in the ways that we infer they will behave when we use the theory to makes inferences about their dynamics. This is the counterpart to the first constraint: an entrance rule for selection theory should not be too weak and let in too much, such that the inferences one makes using the theory are false for a system that meets the conditions stated in the entrance rule.

- An entrance rule for selection theory should be *complete*. It must be possible to follow the entrance rule for selection theory without already knowing anything about what sorts of systems meet the requirements for selection. An entrance rule for selection theory should not presume even a partial understanding of what sorts of systems fit selection theory. An entrance rule that is incomplete is not a real entrance rule for the theory; it is at best a contribution to our understanding of the entrance rule for selection theory, one that requires supplementation with some prior understanding that the reader is presumed already to have. Insofar as Lewontin makes appeal to the notion of population in stating his three principles, his entrance rule remains incomplete (see section 2.5.4).

- An entrance rule for selection theory should not require that, in order to determine that selection theory is deployable upon some systems, the theorist must recognize that the system exhibits what is to be explained by the theory. The danger here is that of circular inference. In order to deploy selection theory to *explain* some phenomenon, we must *infer* that phenomenon. Whatever more determinate features inferences must have

in order to be explanatory, an *explanans* had better provide inferential grounds for its *explanandum*. However, if we have already inferred the applicability of selection theory on the basis of what we are seeking to explain, we will be inferring in a circle if we go on to infer the *explanandum* from the *explanans*. When evolution is made a requirement for selection, a mistake of this sort is made.

- An entrance rule for selection theory should imply that the system picked out has values for all the variables that are deployed in selection theory models. Consider that it should not happen, for instance, that some system might fit the entrance rule for selection theory but fail to be the sort of thing that is found in populations, for it is impossible to deploy selection theory models without considering relative frequencies. The same is true for any variable used in selection theory models. The entrance rule for selection theory triggers the deployment of the whole theoretical apparatus, and it should be possible to follow the rules that characterize the use of that apparatus for every system that meets the requirements for selection. I do not discuss any takes on the requirements for selection below that violate this last constraint, but I mention it because it strikes me as correct, and it's a constraint that my own account of the entrance rule for selection theory fulfills.

In the next few sections, I use the above criteria to evaluate various proposals for the requirements for selection. There are certainly a great many statements of the requirements for selection to be found in both the philosophy and biology literature, so I will concentrate my attention on only some of the most prominent among them. I have chosen my targets because

they say things that indicate they are treating their statements of the requirements for selection as (what I would call) entrance rules for selection theory. But the point is not to clear the field of all rival statements of entrance rules for selection theory before my own proposal is made.

2.2 THE "INTERACTOR" DOES NOT HELP STATE THE REQUIREMENTS FOR SELECTION

In this section, I quickly note that David Hull's notion of "interactor" is not designed to contribute to a statement of the requirements for selection. Hull's notion of replicator can be used to state the requirements for selection, and below I consider Dawkins' use of "replicator" to state the requirements for selection. But Hull's interactor is not even supposed to delineate the applicability of selection theory any more finely than does the notion of replicator on its own.

Hull offers his replicator/interactor pair as a solution to an ambiguity inherent in the phrase, "unit of selection" (Hull 1988, 414). Hull is almost apologetic about his reliance on Dawkins' work, but he sees himself as adding to it, filling out Dawkins' view: "My emphasis on Dawkins is also in part idiosyncratic because it was through my efforts to find out what was missing in Dawkins's analysis that I came up with my own" (Hull 1988, 414). Though Hull thought the concept of interactor was a necessary complement to the replicator, what is critical for our purposes is that his adding of interactors into the picture does nothing to further constrain what sorts of things can undergo selection. I say this because genes, the paradigm replicators, may count simultaneously as replicators and interactors:

> Many other entities may function as interactors—even genes. Genes have "phenotypes." DNA is a double helix that can unwind and replicate itself. In doing so it interacts with its cellular environment. In the beginning, the same entities had to perform both functions necessary for selection. They had to replicate, and they had to interact with their environments in such a way that replication was differential. (Hull 1988, 409)

Because alleles are paradigm replicators, and they can function as interactors, I propose simply to take it that a replicator without any higher-level interactors is a replicator *and* an interactor, and hence that there are no replicators that fail to undergo selection because of the lack of some interactor to figure in the process.

The work of other authors on the units of selection question can equally be set aside as irrelevant to the entrance rule for selection theory. For instance, Sterelny, Smith, and Dickison's discussion of the extended genotype is clearly not intended as an entrance rule for selection theory (1996). Those writers isolate genes and a few other developmental resources as adaptations for the transmission of information down lineages. Since they are treated as adaptations, Sterelny et al. must take the attitude that genes evolved by natural selection to transmit information; accordingly they cannot be requisite for selection to occur.

2.3 RESEMBLANCE AND TYPE SELECTIONIST APPROACHES TO THE REQUIREMENTS FOR SELECTION

The formulations of Darwin's theory found in Lewontin (1970; 1978), Maynard Smith (1983; 1987; 1988; 1991), and Griffiths and Gray (2001; 2004) are reasonably interpreted as putative entrance rules for selection theory. These works share something more in common; they all invoke what I call *resemblance selectionist* stances on the requirements for selection.

I call these stances resemblance selectionists because they pick out large macro-systems, organisms or developmental systems, which are said to undergo selection provided they produce offspring that resemble them (and have other features). In stating the requirements for selection, resemblance selectionists make crucial use of the notion of inheritance to secure the resemblance relationship between parents and offspring. Informally, heredity is typically understood by resemblance selectionists as a matter of resemblance between parents and offspring, a correlation or a covariance in traits or fitness between ancestors and descendants (Lewontin 1970; Brandon 1990, 6; Godfrey Smith 2000, 13; Arnold and Fristrup 1982, 116; Wimsatt 1980, 143). Heredity is also often presented as a formal notion, often in connection with representations of selection that involve the Price equation (e.g., Okasha 2006). The notion of heritability at work there is the *narrow sense of heritability*, namely the quotient of the covariance of ancestor traits with descendant traits and the variance in the ancestor traits:[8]

$$h = \frac{Cov(z', z)}{Var(z)}$$

Systems are said to undergo selection provided that this quantity takes on a positive value (and other conditions obtain).

In contrast to the resemblance selectionist approach, Dawkins offers a *type selectionist* take on the requirements for selection; for Dawkins, the applicability of selection theory is triggered by relatively small things, paradigmatically rival alleles, which institute selection regimes by virtue of producing descendants of the same type as themselves. For Dawkins, descendants must literally be copies of their ancestors, though this especially strong connection between ancestor and descendant is not a necessary component of type selectionism. The key

[8] Heritability sometimes means "heredity in the broad sense" but I ignore that notion, trusting Rice that "heritability in the broad sense plays no role in evolutionary theory" (Rice 2004, 194).

contrast between resemblance selectionism and type selectionism lies here: Resemblance selectionists claim that selection requires, among other things, entities that produce descendants that *resemble* them, ones that inherit at least some traits; type selectionists claim that selection requires, among other things, entities that produce descendants of the *same type* as themselves.

I mean to use the notions of resemblance selectionism and type selectionism as handy labels in what follows. My opposition to resemblance selectionism does not have solely to do with the fact that it posits resemblances between parents and offspring as necessary for selection. Similarly, though my own view is type selectionist, I do not require that descendants be exact copies of their ancestors.

2.4 VERSIONS OF RESEMBLANCE SELECTIONISM

In what remains of this chapter, I will first lay out some resemblance selectionist approaches to the requirements for selection and then criticize these.

2.4.1 The inheritance of fitness as a requirement for selection

Here are the three principles that form the "logical skeleton" of Darwin's argument, according to the first of two articles in which Lewontin discusses the requirements for selection. Together these principles are supposed to embody the principle of evolution by natural selection:

1. Different individuals in a population have different morphologies, physiologies, and behaviors (phenotypic variation).
2. Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
3. There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable) (Lewontin 1970, 1)

Lewontin writes that the logical skeleton of Darwin's argument "turns out to be a powerful predictive system for changes at all levels of biological organization" (1970, 1). In treating Darwin's argument as a predictive system, Lewontin is essentially treating it as a theory.

Lewontin regards Darwin's argument as providing a means of generating predictions about the evolution of biological systems, and providing predictions is the sort of thing that theories can in general be expected to do. The main point of Lewontin's article is to expose the generality of the three principles. Provided the principles apply to some population, Lewontin writes, "the population [will] evolve whether the correlation between parent and offspring arose from Mendelian, cytoplasmic, or cultural inheritance" (1970, 1). Lewontin goes on to consider a number of different non-conventional sorts of systems over which might evolve by natural selection. Presumably, because the same "skeletal" features are present in every application of the theory, Lewontin seems to be offering these principles as an account of the features shared by every system that evolves by natural selection. Other writers have offered similar statements to Lewontin's; Sober writes: "Natural selection impinges on a set of *objects* if there are (heritable) differences in fitness between them" (1984, 216).

2.4.2 The inheritance of traits as a requirement for selection

Lewontin's (1970) statement of the requirements for selection should be contrasted with a similar statement in a later work (1978), in which that author once again states that the theory of evolution by natural selection rests on three principles:

> Different individuals within a species differ from one another in physiology, morphology, and behavior (the principle of variation); the variation is in some way heritable, so that on the average offspring resemble their parents more than they resemble other individuals (the principle of heredity); different variants leave different numbers of offspring either immediately or in remote generations (the principle of natural selection). (1978, 220)

This later statement is in many ways similar to the earlier one, but there are some crucial differences in the formulations. For instance, according to the later article, the different

individuals must be within the same species. Most crucially for our purposes here is how, in the later formulation, it is phenotypic variations, rather than fitness, that must be inherited.

The idea that traits must be inherited (among other things) for selection to occur has had broad appeal among philosophers of biology. Here is a sample of how others have understood Lewontin's principles. Citing Lewontin, Brandon writes that variation, inheritance, and differential reproductive success are necessary conditions for the evolutionary process to work (1990, 6-9). Godfrey-Smith claims "evolution requires a population in which there is variation in phenotype, differential reproduction on the basis of phenotype, and heredity of the traits associated with differential reproduction (2000, 13). Arnold and Fristrup treat their versions of Lewontin's principles as the "conditions necessary for the operation of natural selection" (1982, 116). Wimsatt has written of Lewontin principles that "where (and while) these three principles hold, evolutionary change will occur" (1980, 143). While some of these authors offer variations on Lewontin's three principles, none of the variations is such that the alternative formulation on offer allows its advocate to escape the criticisms below.

2.4.3 Okasha's requirements for selection

Okasha (2006) offers an interesting perspective on the requirements for selection, one motivated by an investigation of how evolution is represented using the Price Equation. While he writes early on that "Lewontin's formulation seems to capture the essence of the Darwinian process very neatly" (Okasha 2006, 13), Okasha later requires that the relative fitness of ancestors covary with the values for traits among descendants rather than ancestors. The point arises due to the non-transitivity of covariance:

> It is possible for x [ancestor trait values] to covary with y [ancestor offspring production], and y [ancestor offspring production] to covary with z [descendant trait values], but x not to covary with z. So even if character and fitness covary, that is, if Cov $(x, y) \neq 0$, it

> does *not* follow that Cov (x, z); and the latter is the fundamental condition that must be satisfied if selection is to lead to evolutionary change. (Okasha 2006, 37)

Ultimately, however, Okasha considers a case in which the "change due to selection" is exactly counterbalanced by "transmission bias," as those quantities are represented in the Price Equation. Effectively, the case is one in which

$$h \bullet \mathrm{Cov}(w,z) = -\mathrm{E}(\Delta z)$$

where h is heritability, w is relative fitness, z is ancestor trait values, E is the expectation operator, and Δz is change in trait value between ancestors and descendants. Okasha writes of the circumstance in which selection is counterbalanced by transmission by saying that "this is *not* an objection to the sufficiency of the Lewontin conditions; in such a circumstance selection *does* still produce an evolutionary response, it is just exactly offset by transmission bias" (Okasha 2006, 39). Effectively, then, Okasha is treating it a sufficient condition for selection that

$$h \bullet \mathrm{Cov}(w,z) \neq 0.$$

Provided that the above condition holds, Okasha will allow that some population is undergoing selection, even if the relative frequency of types within it, or values for traits within it, remains the same over generations. Although this is an interesting perspective on the requirements for selection, the counterexamples considered in Section 2.5 below are still counterexamples to this view, since they are cases of selection in which heritability is zero.

2.4.4 Maynard Smith's formulation of the requirements for selection

Similar to Lewontin's (1970) characterization of the requirements for selection is the formulation of Darwin's theory we find in Maynard Smith's writings. While this formulation is played down in his most recent work where replicators take center stage (e.g., Maynard Smith and Szathmáry 1995; see also Maynard Smith 1987, 124), Maynard Smith put forward a characterization of selection in earlier writings that is worth quoting:

> Darwin's theory can be summarized as follows: Suppose there is a population of entities with the three properties of multiplication (one can give rise to two), variation (not all entities are alike), and heredity (like usually begets like in the multiplication process), and suppose also that some of the differences between entities influence their likelihood of surviving and reproducing (i.e., their "fitness"). Such a population will change in time—"it will evolve." Further, the individual entities will come to possess traits that increase their likelihood of survival and reproduction—i.e., "adaptations." This statement, I think, is not a testable scientific theory but follows necessarily from the original assumptions (including the assumption that there is a continual supply of new variations, some of which increase fitness). (Maynard Smith 1991, 27; see also Maynard Smith 1988, 222; and Maynard-Smith 1983, 316)

Note that, like Lewontin, Maynard Smith takes himself to be characterizing a theory, specifically Darwin's theory. Once again, evolution is made a feature of the formulation, though it is presented as consequence of variation, multiplication, and heredity when an additional supposition is made, there are differences among population members that *influence* their likelihood of survival and reproduction. Like Lewontin, Maynard Smith offers no argument that his characterization is an adequate formulation of Darwin's theory, but he is also explicit about why he does not do so. Maynard Smith regards the statement as a logical truth (1991, 27).

2.5 CRITICISMS OF RESEMBLANCE SELECTIONISM

In this section, I criticize the various resemblance selectionist approaches just considered on three grounds. Whether it is phenotypic traits or fitness that putatively must heritable for selection to occur, either way, the notion of inheritance is simply inadequate to formulate the requirements for selection. Moreover, the notion of "population" at work in resemblance selectionism is left dangerously unspecified. Finally, the macro-systems, the things which putatively must exhibit ancestor-descendant resemblance relationships, are either left undefined and indeterminate, or are defined in a manner that makes it impossible for them to figure in a statement of the requirements for selection, as is the case with developmental systems theory.

2.5.1 Inheritance of fitness is not a requirement for selection: selection for polymorphisms

In the 1970 article in which Lewontin proposed that the inheritance of fitness was a requirement for selection, Lewontin was explicit that he is not seeking to state the requirements for selection but rather the requirements for *evolution by natural selection*. That systems undergoing selection must evolve is a consequence of Lewontin's criterion that selection requires the inheritance of fitness. Lewontin recognizes this; he explicitly excludes a case of overdominance, in which heterozygotes outcompete homozygotes leading the population to rest at a stable polymorphism, from the purview of Darwin's theory (1970, 1).

From the perspective of the current work, Lewontin's exclusion seems strange. For one thing, the development of models of selection that yield polymorphisms is one of the most active fields of research in population genetics modeling of selection (Levins 2004, 22). What's more, the various population genetics models that yield persistent polymorphisms may be indistinguishable from those that yield fixation of one type or another in the population except insofar as they involve different values for their constituent variables.

The best known model that yields a stable polymorphism is a diploid selection model with overdominance, in which heterozygotes have the greatest viability. Alter nothing in such a model but a single fitness coefficient, such that the heterozygotes have a fitness value between those of the two homozygotes, and the model becomes a model of directional selection, yielding evolutionary change toward fixation. Surely it is a weird categorization that places in different theories these two models that differ only in the value for one their variables. Indeed, things are worse than this, for Lewontin and his followers will be forced to regard a population which is evolving *toward* a stable equilibrium as undergoing selection, but must then cease to regard it as undergoing selection once it has already attained the equilibrium point and remains stuck there. Since the same model yields evolution when the population is away from its intermediate equilibrium frequency, but stasis once that equilibrium is reached, Lewontin is in the awkward position that he must regard a population governed by a single model over an extended time period as undergoing selection at early times when evolving toward a stable equilibrium but not at later times once it has reached it.

Even worse still for the resemblance selectionist view that makes evolution requisite for selection, some sustained polymorphisms may also be adaptations, the sorts of things we need "Darwinian theory" to explain, or at least make explicable. Here is a compelling example of a polymorphism that is an adaptation. The bacteria, *Haemophilus influenzae*, develop appendages called fimbriae that allow them to bind to host mucous and epithelial cells and are therefore useful in colonization. The fimbriae become deleterious, however, once the host has been infected and the pathogens have infiltrated the spinal cord fluid, because the fimbriae make the invaders easier to detect by the host's immune system at that point in their lifecycle.

The expression of fimbriae is controlled by what Moxon et al. call discriminate stochastic mechanisms of gene expression, also known as contingency genes (Moxon et al. 1994, 23-27). Contingency genes control a wide range of pathogen features involving evasion of immune system responses. These genes mutate often enough that descendants are likely to express different phenotypes than did their ancestors, allowing the bacteria to have offspring displaying different phenotypes, at least some which will be well-suited to develop in an environment that their ancestors invaded, but which they did not initially inhabit.

This sustained polymorphism makes for a compelling case of an adaptation, one that not only evolved by selection, but which also *persists* by selection. The ongoing presence of this polymorphism can be made explicable by appeal to how it is conducive to the reproduction of lineages of bacteria that have it:

> Genomes can respond to unanticipated challenges. This is because natural selection has selected genomes that best survived unexpected challenges in the past. For many bacteria species, this flexibility results from the presence in the genome of hypermutable contingency loci, which provide a repertoire of variation, allowing the population to adapt rapidly in the face of unpredictable contingencies, such as changes in the host environment. (Bayliss, Field, and Moxon 2001, 657)

If we imagine competition between lineages of *H. influenzae* with the tendency to mutate at contingency loci and other rival lineages of bacteria that do not tend to mutate at these loci and are either fixed for fimbriae expression or the lack thereof, we can see how the hypermutants would outcompete their rivals. Hypermutants with fimbriae that aided them in gaining access to a host's spinal fluid will produce a significant number of descendants who lack fimbriae and are harder to spot within the fluid. These will produce descendants who go on to colonize other hosts aided by the fimbriae that they develop owing to yet another mutation. In contrast, one stage of the lifecycle or the other will institute severe selection against any lineage fixed either for fimbriae or the lack of fimbriae.

While Lewontin's principles make evolution requisite for the application of "Darwin's theory," Darwin himself offered selectionist explanations of polymorphisms too. His study of heterostyled plants is a case in point (1989[1884])[9]. Plants exhibiting heterostyly develop two, or sometimes three, different forms of flower whose reproductive organisms vary in a number of ways, principally length. Some plants exhibit different forms of flower on the same plant, while some are dimorphic and trimorphic, with only one sort of flower per plant. Darwin interpreted the flower variations as conducive to intercrossing, which he thought was beneficial, at least for many organisms. All the plants exhibiting heterostyly that Darwin considers are pollinated by insects, and the varying lengths of the different sex organs of the plants cause the organs to brush against different parts of insect bodies (Darwin 1989[1884], 69; 105). Thus, pollen from the anthers of flowers with long stamens will be deposited on the stigma of flowers with long pistils, and *mutatis mutandis* for the reproductive organs of converse lengths. Darwin thinks it clear that heterostyly is an adaptation:

[9] Thanks to Prof. Lennox for pointing me to this example in Darwin's work.

> The benefit which heterostyled dimorphic plants derive from the existence of the two forms is sufficiently obvious, namely, the intercrossing of distinct plants being thus ensured. Nothing can be better adapted for this end than the relative positions of the anthers and stigmas in the two forms" (1989[1884], 22)

The fact that heterostyly is scattered among diverse groups of plants is further evidence that it is the product of evolution by natural selection:

> It has been shown that heterostyled plants occur in fourteen natural families, dispersed throughout the whole vegetable kingdom, and that even with the family of the *Rubiaceae* they are dispersed in eight of the tribes. We may therefore conclude that this structure has been acquired by various plants independently of inheritance from a common progenitor, and that it can be acquired without any great difficulty – that is, without any very unusual combination of circumstances" (Darwin 1989[1884], 189).

While Darwin is confident that heterostyly is an adaptation designed to promote cross fertilization, he does struggle a little bit to explain how heterostyly could have evolved by natural selection. He imagines that heterostyled species initially exhibited variability in their reproductive organs and, though already fertilized by insects, could have gained from even greater cross-fertilization. Darwin imagines that the "law of compensation," according to which enhancements in some plant structures are paired with reductions elsewhere, could have facilitated the division of a continuously varying population of into a dimorphic one by associating short pistils with long stamens and vice-versa. Selection could then have cemented the dimorphism, as those plants that demonstrated it would intercross more freely and hence spread in the population.

The chief difficulty that Darwin's explanation faces is that he cannot account for why heterostyly should evolve as a mechanism to facilitate intercrossing rather than alternative mechanisms, ones that do not have the consequence that the plant cannot reproduce with half of its conspecifics, a liability for heterostyled plants. Indeed, heterostyly is but one of several forms

of polymorphism that are adaptations for intercrossing; Darwin also regards as adapted for intercrossing plants that display self-sterility and plants that display dichogamy, in which the timing of maturation of the male and female sex organs differs (1989[1884], 7; 190-191). Darwin simply supposes that heterostyled species did not possess the right sort of variations to evolve alternative means of ensuring cross-fertilization:

> It might well happen that our supposed species did not vary in function in the right manner, so as to become either dichogamous or completely self-sterile, or in structure so as to ensure cross-fertilization. If it had thus varied, it would never have been rendered heterostyled, as this state would then have been superfluous." (1989[1884], 189).

Whatever the fate of Darwin's demonstration of the explicability of heterostyly, it remains the case that he regarded his theory as one that could explain the persistence of polymorphisms in populations. So it is critical that one not make the same mistake as did Lewontin and his many followers of presenting Darwin's theory as one for which evolution is a necessary condition.

I think that heterostyly and the fimbriae polymorphism of the bacteria studied by Moxon and others make for compelling cases of adaptations, adaptive polymorphisms whose designed-appearing character invites demystification by selection theory. However, it is difficult to determine whether they are adaptations because the term "adaptation" is vague. As it is being used here, the term adaptation refers to designed-appearing characteristics, and it is hard to judge what exactly counts as "designed-appearing" in borderline cases. "Adaptation" can be made rigorous by attaching it explicitly to selection theory in this sort of way: adaptations are what are produced as a result of selection. That proposal, however, makes it impossible to use selection theory to make adaptations explicable, since selection produces adaptation by definition. So whether or not we should count as adaptations those characteristics that persist as polymorphisms depends on whether or not they are sufficiently designed-appearing, a difficult judgment to make.

However, the very indeterminacy of the term "adaptation," when used in its non-rigorous sense as "designed-appearing character," stands as a further reason to think that refusing to countenance polymorphisms as adaptations is a bad idea. It is simply *unclear* what traits exactly should count as adaptations. We are better off, then, stating the conditions necessary for the *deployment of selection theory* rather than the conditions necessary for *adaptive evolution*, since selection theory can make explicable adaptations and at least some polymorphisms, whether or not we count the latter as adaptations.

2.5.2 Inheritance of fitness is not a requirement for selection: circularity

In his 1970 article in which Lewontin claims the inheritance of fitness is a requirement for selection, Lewontin could only mean by "fitness" actual reproductive output, rather than what is represented by fitness coefficients in population genetics equations, for the latter may vary even

when the population has reached a stable polymorphism. But Lewontin's requirement that actual reproductive output be heritable creates another set of problems. To determine whether or not fitness is heritable, we must first know whether or not the population will be one in which evolution is ongoing, something that should bar us from inferring (and hence bar us from explaining) that evolution. We have to know whether individuals that tend to produce especially many offspring tend to have descendants that do so too. The difficulty is that there is no way to know that without first checking how many descendants the descendants go on to produce.

Evolution can reach stable polymorphisms such that selection leads one generation of individuals to produce especially many offspring while their descendants, now more frequent, do not do so. So we cannot infer from the fact that especially fertile parents have offspring that share their genes, or even are identical to them in every respect, to the fact that the offspring they produce will be especially fertile. So by stipulating that there must be heritability of fitness for evolution by natural selection, Lewontin is effectively requiring that the theorist first recognize that the system of interest is one whose dynamics are already understood, prior to recognizing that selection theory is deployable over it. Other authors explicitly require evolution as a necessary condition for selection (Arnold and Fristrup 1982, 116; Sober 1984, 216), and these writers follow Lewontin in violating constraint number four. If Darwin's theory is to function to explain the evolution of some system, or more generally the dynamics of some system, it must be possible to *infer* that evolution is taking place within it. On pain of circular reasoning, we must not use Lewontin's requirements for selection as a basis for inferring that selection theory is applicable to some system, at least if we plan to use selection theory to do any explanatory work.

2.5.3 Inheritance of traits is not a requirement for selection

While the inheritance of fitness is not a requirement for selection, neither is the heritability of phenotype. This is true as long as one conceives of inheritance in one of the standard resemblance selectionist fashions, as a relationship between the traits of ancestor and descendant macro systems, paradigmatically organisms. Heritability in the narrow sense will be zero if ancestor and descendant trait values do not covary. As the following counterexamples show, the phenotypes of macrosystems need not covary with the phenotypes or fitnesses of their descendants for selection to go on in some population.

It is quite easy to set up a population genetics model in which selection occurs among macro systems whose traits, and indeed whose reproduction rates too, are not inherited, that is, a model in which the organisms are just as likely to resemble the parents of their rivals as they are to resemble their own parents in their phenotypic traits. Consider a variable selection model featuring the random assortment into two equally common sub-environments of diploid individuals differentiated by genetic variation at a single locus exhibiting complete dominance. The selection is symmetrical and paradoxical, such that each individual has the same relative fitness as does its rival in the opposite sub-environment because each individual has the same phenotype as does its rival in the opposite sub-environment.

We can imagine that the phenotype in question is camouflage, imagining prey insects of different tints living in a patchy forest. In one sub-environment, one phenotype is well camouflaged but the other is highly conspicuous, while in the other sub-environment, the opposite is true. A population under selection of this sort will produce a stable polymorphism at equal relative frequency for both variants (Levene 1953, see section 6.3.5.2.3 where I generate such a model using the cause-to-model algorithm developed later on). What's more, population

members are just as likely to resemble the parents of their rivals as their own parents, since they are just as likely to turn up in one sub-environment as the other because of the random assortment of individuals into sub-environments is assumed. So, by using standard population genetics theory, one can easily produce a model in which there is selection, selection that *explains* the dynamics of the population, but in which there is no inheritance of the phenotypic trait of camouflage. Because the fitness values in the model are symmetric, too, there is no inheritance of fitness, either.

As another example, consider a sex-dependent selection model with genetic variation at a single locus that contributes in a paradoxical fashion to a quantitative trait. Assume an intermediate level of dominance, and imagine that one allele is beneficial among males and elevates the value for the quantitative trait, while among females, the allele is deleterious and depresses values for the trait. Just as in the scenario just considered, an allele with this sort of influence will persist in the population at a stable intermediate frequency because of the action of selection (Hedrick 2005, 179; see section 6.2.3.3.2.2 where I generate a sex-dependent selection model of this sort). . Furthermore, since individuals are just as likely to have male offspring as female offspring (let's imagine), the heritability of the quantitative trait the allele impacts will be nil. The heritability of fitness will be nil, too. The male and female homozygotes who take on values for the trait that deviate from the mean are just as likely to have offspring that have values that deviate from the mean in either direction. Their descendants do not inherit their parents' deviations. This is because the parents are just as likely to have homozygote offspring that are males as ones that are females, while their heterozygote offspring are average. Heterozygotes are just as likely to have both sorts of deviant offspring too. Once again, this is a simple scenario involving a simple population genetics model, and it is a strange conception of the requirements for selection that does not treat the scenario as an instance of selection.

Presumably, that some scenario can be captured using textbook classical population genetics theory through attributions of relative fitness values to variant types is sufficient grounds for regarding the scenario as an instance of selection. Indeed, this sort of sex-dependent selection model has been applied to human populations in an attempt to explain male homosexuality as the result of genetic variations that persist in human populations because of their sex-dependent impact on fertility. Male homosexuality is posited to be the result of a

variant allele that increases female fecundity but, because it causes homosexuality in males, lowers male fecundity (Gavrilets and Rice 2006). At equilibrium frequencies, such a gene will not produce heritable deviations in fecundity because especially fecund females will be just as likely to produce especially infertile males as especially fecund females.

The above sorts of counterexamples could be multiplied. Generally, parents may fail to resemble their offspring because the alleles that they pass on to their descendants interact with causal contexts that are distinct from the causal contexts with which the alleles interacted among the parents. Phenotypes are the causal product of more than just alleles; they are the causal products of alleles and other causal factors that interact with alleles. Set up the impact of the other causal factors in just the right way, and one can produce a situation in which the phenotypes of offspring do not resemble those of the parents and equally cases in which the fitness of offspring does not covary with that of their parents, fitness being just another phenotype.

Surely, classical population genetics must be understood as one of the central ways in which selection is modeled: Population genetics models of the sort I put forward as counterexamples to the resemblance selectionist view include differences in relative fitness between individuals, and they are the sorts of models that are discussed in population genetics textbooks in chapters dealing with natural selection, and they are the sorts of models used to make explicable natural phenomena as the result of selection. Accordingly, such models can hardly be treated as ones that do not fall under purview of selection theory. So the scenarios should be regarded as counterexamples to the resemblance selectionist view that inheritance is a necessary condition for selection. They demonstrate that neither phenotypes nor fitness must be heritable in the formal or informal sense for selection to occur.

In each of the counterexamples above, I set up the causal impact of a single non-allelic factor (environment or sex) in such a way that descendants do not resemble parents. Presumably there are more complex ways in which this same lack of resemblance can occur, ones involving more than two alleles, multilocus selection, more complex forms of variable selection, sex-dependent selection, antagonistic pleiotropy, or combinations of these.

Because whether or not a system exhibits inheritance of traits and whether it exhibits inheritance of fitness are independent, we can generate models similar to the ones above, but in which different sorts of dynamics are possible. For instance, we can imagine an allele that spreads to fixation despite producing a trait that is not heritable. Imagine an allele that impacts height, making males taller and females shorter. Such an allele will not induce heritable differences in height in a population with an even sex ratio. But if males profit from their extra stature while females do not suffer from being short, then the allele will spread to fixation in short order.

A further difficulty with the resemblance selectionist view of selection is that it is ambiguous: whether we judge selection to be occurring or not will depend on what trait we use to assess whether the population exhibits inheritance. To see this, let's return to the first counterexample scenario, and consider the pair of alleles at a single locus that contribute to fecundity in a paradoxical and sex-dependent fashion. Let's imagine further that the alleles do so by way of a causal intermediary, say a hormone, quantitative values for which are independent of sex. One allele produces elevated levels of the hormone in both sexes. Elevated levels of the hormone in turn have paradoxical effects in females and males, increasing the fertility of the former while reducing that of the latter.

In this sort of scenario, if we consider fecundity to be the trait of interest, the trait will not be heritable for populations fitting our sex-dependent selection model. But if we count the *hormone* as the phenotype of interest, the trait of elevated or depressed hormone levels *will* be heritable, for population members with elevated levels of the hormone will tend to have offspring with elevated levels too. By trading in tint for camouflage in the spatially variable selection model considered earlier, the same transformation from a system that fails to meet the requirements for selection into one that does is effected, since tint can be inherited while camouflage is not.

The dynamics of the fertility of a population under sex-dependent selection are the same whether we assess the population as undergoing selection or not. The dynamics of the hormone that impacts fertility are the same, too, no matter what our assessments, as are the dynamics of camouflage and tint in the variable selection scenario considered. The dynamics of all these traits, along with the alleles that causally contribute to them, are the result of selection; we can model the active alleles that contribute to the development of these traits using classical population genetics models of selection. But if the selection dynamics of the alleles and traits do not depend upon what we identify as the trait of interest, while whether or not the population exhibits the inheritance of traits does depend on the trait upon which we choose to focus, then inheritance cannot effectively contribute to a statement of the requirements for selection.

That there are heritable developmentally primitive traits lying beneath every more developmentally remote non-heritable trait might seem to offer an escape route for the resemblance selectionist. It appears as though the resemblance selectionist can ward off the counter-examples I propose provided that individuals applying selection theory are required, somehow, to focus on the heritable traits that are the developmental antecedents of the non-heritable ones when assessing whether a system is undergoing selection. There are two difficulties with this escape plan.

First, the resemblance selectionist will need some non-circular way of making sure that the right traits are considered when it is assessed whether some system meets the requirements for selection. The counter-examples make untenable the possibility of deciding whether the dynamics of some *trait* in a population is undergoing selection on the basis of its heritability. What the escape route offers to the resemblance selectionist is the possibility of judging whether some *system* is undergoing selection on the basis of inheritance. If the resemblance selectionist is to maintain her view, she will have to offer some recipe for picking out traits when using inheritance of traits to assess whether a system is undergoing selection, a recipe that will always dictate that a heritable trait is chosen whenever the system actually is undergoing selection. Without such a recipe, the resemblance selectionist stance on the requirements for selection cannot be trusted. So, at best, the resemblance selectionist account of the requirements for selection is incomplete and in its current form cannot function as a statement of the requirements for selection.

The second problem with the escape route we are considering is that it leads ultimately to the type selectionist view. To see this, recall how in the sex-dependent selection model we imagined that fertility was not heritable while hormone levels were heritable. Hormone levels

causally interact with sex differences, so even though they are heritable, they do not have a single, sex-independent impact on fertility, which is why fertility is not heritable too. But the same dynamics for fertility will result if we imagine that hormone levels are themselves the result of an even more developmentally primitive cause that interacts with sex differences. We could imagine that one of our alleles has a sex-independent impact on a structural feature of the endocrine system. That structural feature then has a sex-dependent impact on hormone production, causing elevated production of the hormone in males and reducing it in females. The hormone in turn has a sex-independent influence on fertility. In this sort of case, fertility will not be heritable, hormone levels will not be heritable, but structural features of the endocrine system will be heritable.

We can equally imagine a population of insects in which neither camouflage nor tint is inherited, again by bringing into the picture more primitive developmental causes. Imagine that the insects' environment features two different food resources, and that the tint of the homozygotes is contingent upon the food they consume. Make the distribution of food resources even, determine what food an insect consumes randomly, and have one type of homozygote come out dark when it consumes the first resource and light when it consumes the second, while the opposite occurs in the other homozygote. In such a case, we have a system in which camouflage is not inherited, tint is not inherited, but divergent developmental responses to food resources are inherited.

All we have done in the above transformations of our earlier cases is to push the causal interaction between the alleles and their causal context to a more developmental primitive stage. Presumably, we could pursue this sort of regress still further. If we do so, the resemblance selectionist who seeks to avoid the counter-examples will have to insist at each step that what

before was the phenotype used to assess whether the system is undergoing selection should no longer be used. At each step, she must tell us to judge whether the system is undergoing selection by assessing the heritability of a more developmentally primitive phenotype. She cannot say, anyway, "look not to fertility, but to hormones," or, "look not to camouflage, but to tint."

There is a firm stopping point to this regress, however. That stopping point occurs at the genetic variations whose developmental impact is contingent upon the interactive contextual causes posited in the models. It will never come out that a population genetics model of selection shows that the trait of bearing a specific allele is not heritable.[10] This suggests that the resemblance selectionist is best off simply making the trait whose heritability should be assessed when deploying her requirements for selection the trait of bearing one or another causally active allele at one or more loci. But in doing so, the resemblance selectionist is effectively undertaking the type selectionist stance, since the inheritance of alleles requires their germ-line replication, the production of copies of the alleles, and vice-versa. Even when the resemblance selectionist *could* pick out some trait other than bearing one allele rather than another as the heritable trait she uses to affirm that selection is going in some system, it is nonetheless always possible for her to check for the inheritance of genetic variations that causally contribute to the trait. So, type selectionism stands as a simpler and more general alternative to resemblance selectionism, at least for the sorts of cases considered above.

[10] This is true, anyway, for mutation rates of less than 0.5.

What parents have in common with their offspring is the alleles they contribute to them, alleles that may then react with environmental parameters in genotype-specific ways. The deployment of classical population genetics at least requires the transmission of *alleles* from parents to offspring, and hence the creation of descendant alleles that are of the same kind as the parental alleles. Even if we generalize to include cases of expanded inheritance, then the deployment of population genetics models requires the transmission of developmental resources that can be treated as alleles are treated in classical population genetics theory.

So the resemblance selectionist perspective that depends upon the inheritance of traits to state the requirements for selection accordingly fails to do so adequately because inheritance of traits is simply *not* a requirement for selection. Only if we regard the alleles themselves as "traits" that are inherited will we avoid the sorts of counter-examples just discussed, but to do so is simply to give up the resemblance selectionist view to undertake the type selectionist one.

2.5.4 Resemblance selectionist requirements for selection are incomplete: the notion of population

The third way noted above in which statements of the requirements for selection can fail to pass muster as entrance rules for selection theory is to be incomplete. Ideally, an entrance rule for selection theory will tell you each and every feature requisite for the applicability of the theory. A candidate entrance rule is incomplete if it cannot be deployed effectively unless one already understands at least *some* of the features a system must have to undergo selection. Any statement of the requirements for selection that leaves the notion of "population" unspecified, such as Lewontin's initial statement, is incomplete in just this way.

As stressed in the introduction, populations can be delimited in correct and incorrect ways for the purposes of doing selection theory. One need only imagine wholly wrong-headed

putative populations to see this, such as ones that bring together individuals living great distances apart and separated by insurmountable geographical barriers, or ones involving organisms that hardy interact with each other despite living in the same region (see discussion in Damuth 1985).

If there are incorrect ways for delimiting populations for deploying selection theory, there must be rules, or at least norms, for delimiting them in a suitable fashion for deploying selection theory. Lewontin is requiring of his readership that they understand such rules or norms, for were his three principles applied to an unsuitable population, the usual consequences of natural selection would not result. The interpretation of "population" made on the part of Lewontin's readership is making a substantial contribution to Lewontin's principles; the principles only have a shot at correctly picking out systems over which selection theory may be deployed when population is interpreted in the right way. So by using the term "population" in his statement of the requirements for selection, Lewontin is effectively requiring that his readership already know something about the requirements for selection.

While Lewontin does not mention the term "population" in the second set of principles (1978), he there uses a taxonomic category, "species," to delimit populations. This is a step backward, for there is no reason to restrict the deployment of selection theory to populations composed of conspecifics. Furthermore, members of a single species need not be anywhere near each other on Earth, and the dynamics of such disjoint populations need not have anything to do with one another.

2.5.4.1 An objection. Lewontin might retort that any statement of the requirements for selection, indeed any statement of the requirements for anything, is put forward under the presumption that one's audience already knows how the bits of vocabulary used to state the requirements should be understood. If failing to state the use of each bit of vocabulary used in formulating a set of

requirements is a mistake, then such mistakes are inevitable. There is a clear regress here, one in which rules are required for the use of terms that state rules, and rules required for the use of the terms that state the rules for stating the rules, and so on. This sort of regress has nothing special to do with selection theory, and it would seem that I am making the impossible demand upon Lewontin that he halt it.

The sort of retort just considered is unavailable to Lewontin. It would be available were the notion of "population" at work in selection theory not a specialized one proper to selection theory alone. However, as I have already pointed out, the use we make of the notion of population in selection theory is indeed a specialized one, proper to selection theory alone. Populations circumscribed for the purposes of taxonomy make poor populations over which to deploy selection theory. So do the ones used in ecology and so do ones from systematics (these points are developed further in the following chapter). Indeed, in his 1970 article, Lewontin himself recognizes that members of a population need not be conspecific organisms, or even organisms at all, for selection theory to be deployed over them. He discusses all sorts of non-standard cases of selection, including selection among molecules, organelles, and groups. No other science considers groupings of such wide variety of entities such that selection theorists could simply borrow the notion of population at work in that science and use it for their own purposes. Lisa Gannett has argued that biologists group entities into populations in different ways for different theoretical and explanatory purposes (2003); she gives good reasons for us not to assume that a population that is circumscribed by biologists for a given purpose will serve well as a population over which to deploy the apparatus of selection theory.

That "population" in selection theory is a specialized notion, one proper to selection theory alone, means that implicit knowledge of how to group entities into populations for the

purposes of doing selection theory must be implicit knowledge about how to deploy selection theory. Because Lewontin assumes such knowledge on the part of his audience, his statement of the requirements for selection is incomplete. Knowledge of how to group entities into populations for the purposes of deploying selection theory is not the sort of knowledge one could come by in any other way except through an understanding of the theory itself (though not necessarily an explicit understanding that would take the form of a definition).

Contrast "different," another bit of vocabulary from Lewontin's principles. In using the term "different" to state his three principles, Lewontin is presuming that his readership understands what "different" means; he does not define the term. But an understanding of "different" is something one could acquire without having any understanding of selection, evolution, or anything else even remotely Darwinian in provenance. "Population suitable for selection theory" is not like this precisely because an understanding of how to group entities into populations gleaned in other circumstances, say, from an understanding of theories in systematics or ecology, does not translate over to the case of selection theory.

So by presuming that his readership understands what "population" means, Lewontin is in effect presuming that his readerships understands at least something about the requirements for selection. This makes Lewontin's principles incomplete when considered as an entrance rule for selection theory. The point is a delicate one, because my own proposal for how to state the requirements for selection does not include a description of how populations should be circumscribed. I am not susceptible to the criticism, however, because I do not deploy the notion of "population" in my definition. It is thus possible for me to regard some system as one over which selection theory is deployable, but yet for which the value of its population size remains a mystery. I say how to group the sorts of entities over which selection theory is to be deployed in to populations (chapter 4) after having said how to pick them out (chapter 3).

2.5.5 Resemblance selectionist requirements for selection are incomplete: macro-systems in selection theory

The last argument showed that statements of the requirements for selection that leave the notion of "population" unspecified are incomplete; they presume a partial understanding of selection theory on the part of those who would deploy them. A similar difficulty emerges when we ask of what sorts of things the populations must be composed. The traditional response to this question might well be organisms, but that response is worrisomely narrow. Classical population genetics models include relative frequency terms for gametes, zygotes, and mating pairs, as well as haploid organisms. Furthermore, the organisms only response rules out any non-traditional deployments of selection theory over such things as colonial organisms, such as Portuguese man-o-wars, or fragments of culture. Moreover, as noted earlier, resemblance selectionists such as Lewontin often explicitly seek generality in their understanding of selection and consider legitimate the application of the theory over populations of entities that are not organisms.

The problem posed by saying exactly what sorts of macrosystems must exhibit inheritance (and other things) to undergo selection has largely gone unrecognized by disputants over the requirements for selection. Griffiths and Gray, two prominent exponents of developmental systems theory, have sought to answer this question because they seek a statement of the requirements for selection that is consonant with the broader tenets of their developmental systems perspective, a novel approach to biological theorizing quite generally. Griffiths and Gray seek (among other things) to generalize the theory of selection by arguing that it should be deployed over developmental systems, a novel type of macro-system that they define. Though they are not the only authors working in developmental systems theory (DST), they have made the greatest effort to show how Darwinism works within the DST viewpoint. So in what follows, I treat their views as constitutive of DST, though there are more workers in the DST tradition than just them.

Inspired by Lewontin, Griffiths and Gray state the requirements for selection in this way: "natural selection occurs because individuals vary, some of these variations are linked to differences in fitness, and some of those variants are heritable" (Griffiths and Gray 2004, 1). Just like those of Lewontin, Griffiths and Gray's principles at least *seem* designed to function as an entrance rule for selection theory. Griffiths and Gray write that "natural selection occurs *because* individuals vary, some of these variations are linked to differences in fitness, and some of those variants are heritable" (Griffiths and Gray 2004, 1; my italics). There is clearly an inferential connection between the three requirements and the occurrence of natural selection, the latter happens *because* the former are fulfilled.

The first thing to note about the Griffiths and Gray formulation of the Lewontin principles is that it does not make the mistake made by most other resemblance selectionists of

tying selection to evolution. Another difference between the DST approach and traditional resemblance selectionist ones is that Griffiths and Gray claim that selection theory is to be deployed over *evolutionary developmental systems* (henceforth just developmental systems), which are defined this way: "the developmental system contains all those features which reliably recur in each generation and which help to reconstruct the normal life cycle of the evolving lineage" (Griffiths and Gray 2001, 207). So the sorts of things that exhibit the features of heritability, variation, and fitness differences in the DST account are evolutionary developmental systems, rather than organisms or other sorts of macro-systems.

Note, too, that Griffiths and Gray define the notion of inheritance at work in their requirements for selection in an unusual way: inheritance is the reliable reproduction of resources down lineages (Griffiths and Gray 2001, 214). The DST notion of inheritance is designed to be deliberately accommodating, allowing all sorts of features of developmental systems to count as heritable ones, not just genetic variations. What's most important about it for our purposes is that it is stated in such a fashion that the counterexamples to the standard resemblance selectionist view that selection requires inheritance that I offered above do not work against Griffiths and Gray. For these writers, inheritance is a matter of the transmission of *resources*, not the recurrence of traits in ancestors and descendants, or covariance in values for trait variables between parents and offspring. Though they would be surprised to hear it, Griffiths and Gray are type selectionists like Dawkins and myself, not resemblance selectionists like Lewontin.

To see this, recall that the difficulty with using inheritance in the standard fashion to state the requirements for selection was that more than just the developmental resources whose dynamics are governed by selection causally contribute to the development of traits. These other

causes can be set up such that even though developmental resources have their dynamics governed by selection, their causal effects, phenotypes, are not heritable. The developmental resources we considered initially were alleles, but the point holds more generally. So, if one makes inheritance a matter of passing on alleles and other developmental resources, as do Griffiths and Gray, then the above counterexamples are evaded. Accordingly, Griffiths and Gray can correctly construe cases of paradoxical symmetrical variable selection and cases of paradoxical symmetrical sex-dependent selection as cases of selection, since in those circumstances, developmental resources (alleles) that matter to fitness are being reliably reproduced down lineages.

Still, there are other ways in which Griffiths and Gray's suggestion for stating the requirements for selection runs into difficulties, ones that are proper to their perspective. Griffiths and Gray violate the fourth constraint from section 2.1; they do not state a complete account of the requirements of selection. One would have to already have a grip on how to deploy selection theory in order to assess whether their requirements hold over some system. Specifically, there is no way to get a grip on what developmental systems are without a prior understanding of selection theory. As we will see, the theorist must know how to deploy selection theory already before she can even make sense of what Griffiths and Gray postulate as the requirements for the theory, for she must understand selection theory to understand what developmental systems *are*. Accordingly, she cannot use Griffiths and Gray's requirements to determine the applicability of the theory, for the requirements cannot state anything she does not already have to know.

To make good on their novel theoretical term "developmental system," Griffiths and Gray must explain two things, how developmental are to be delineated, and how they are to be

individuated. The explanations Griffiths and Gary offer for how the delineation and individuation of developmental systems are to be performed are what undermine their statement of the requirements for selection, for we must already understand selection theory in order to understand these accounts. An explanation of how developmental systems are to be *delineated* consists in saying how to tell apart what sorts of features are to be included in evolutionary developmental systems and what sorts of features must be excluded. Criteria of *individuation* will allow one to tell whether multiple traits and features that each lie within *some* developmental system also belong to a *single* developmental system, rather than belonging to multiple distinct developmental systems.

It is easiest to show that developmental systems cannot be individuated in DST without recourse to an understanding of selection theory because Griffiths and Gray overtly require theorists to understand natural selection in order to individuate developmental systems. Griffiths and Gray discuss the problem of delineation separately from the problem of individuation; the problem of delineation is discussed at length in two early pieces of Griffiths and Gray (1994; 1997), whereas Griffiths and Gray's most recent take on the problem of individuation can be found in a later work (2001, 209-214).

2.5.5.1 Individuating developmental systems. To deploy the requirements for selection theory that Griffiths and Gray offer, we will need to be able to individuate developmental systems. Individuating biological systems is a problem not just for DST but for biologists generally. There are several well-known examples of systems that are hard to individuate, such as slime molds, colonial organisms, and tight symbioses among organisms in different lineages. One promising route to making individuation judgments in such cases is through selection theory. The cost of taking this route, however, is that one cannot use principles for individuation that are arrived at

in this way as inputs to selection theory without engaging in circular reasoning, so requirements for selection that function as requirements for the deployment of selection theory must not make reference to individuals that are picked out in ways that depend upon an understanding of selection theory.

Griffiths and Gray tackle the problem of individuation in a manner similar to that employed by Sober and Wilson. Griffiths and Gray write, "an individual is a system in which the parts form a trait group with respect to most future evolutionary processes" (2001, 213). Following Sober and Wilson (1994), trait groups are in turn defined as "a set of organisms relative to which some adaptation is, in economic terms, a public good" (Griffiths and Gray 2001, 210-11). Notice the use of the notions of adaptation and future evolutionary processes. Griffiths and Gray further assert, "an adaptation is anything that results from natural selection" (2001, 209). This last definition completes the circle: ultimately one must be able to tell what results from natural selection, and hence be able to understand the theory of natural selection, in order to individuate the systems over which the theory is supposed to be deployed. Presumably, however, if one already understands the theory of natural selection, one already knows over what sorts of systems the theory is to be deployed. The Griffiths and Gray requirements for selection cannot be deployed in a non-circular manner because one cannot individuate developmental systems without already knowing what will go on in "most future evolutionary processes" and what counts as an adaptation produced by natural selection, things that one cannot understand without first understanding selection theory and hence understanding over what systems it may be deployed.

2.5.5.2 Delineating developmental systems. Not only is the Griffiths and Gray statement of the requirements for selection circular because of how they individuate developmental systems, but

it is also circular owing to how they delineate them. To successfully delineate developmental systems, one must be able to tell what sorts of things fall within developmental systems and what sorts of things fall without. It is more difficult to show the development systems cannot be delineated without recourse to the use of selection theory because doing so requires that I argue that one putative means that advocates of DST use to delineate the developmental systems does not work. Griffiths and Gray have two putative means to delineate developmental systems, one of which deploys theoretical language from selection theory and one that does not do so. The delineation criterion that does not deploy theoretical language from selection theory fails to perform its delineation function, while the strategy that involves theoretical language from selection theory embroils its user in the same sort of circularity as was exposed in the previous section.

The pressure on advocates of DST to offer rigorous means of delineating developmental systems has its source in Kim Sterelny's "Elvis Presley Problem," a problem Sterelny developed in correspondence with Griffiths and Gray:

> Elvis Presley is part of my [Sterelny's] developmental system, being as he was causally relevant to the development of my musical sensibilities, such as they are. Yet surely there is no system, no sequence, no biologically meaningful unit, that includes me and Elvis. (Griffiths and Gray 1994, 286)

Advocates of DST claim that they have the Elvis Presley problem under control; they are interested only in *heritable* developmental interactions:

> The theory is interested in those developmental resources whose presence in each generation is responsible for the characteristics that are stably replicated in that lineage. For example, we might contrast two influences on a newborn bird. The interaction between the newborn bird and the song of its species, which occurs in each generation and helps explain how the characteristic song is produced, is part of the bird's developmental process. The interaction between the newborn bird and the noise that ruptures its eardrums plays no such role, and is not part of the process. (Griffiths and Gray 1994, 296)

Griffiths and Gray also offer other examples of features that are not produced by characteristic interactions of developmental systems, such as human autism (1997, 476). Indeed, the study of developmental abnormalities requires them to delineate another sort of developmental system entirely (Griffiths and Gray 1994, 287; 1997, 476).

The DST notion of inheritance, however, cannot be successfully used to distinguish between what can and cannot make it into the developmental system or developmental process. The difficulty comes from polymorphisms: while every peculiarity of the development of every individual within a lineage cannot be included as part of the developmental system, there are certainly cases where different developmental interactions should be included within the same developmental system. These include cases where developmental outcomes and processes are only slightly different among descendants than among their ancestors. More interesting are cases in which two clearly distinguishable and exclusive traits should both count as features of single type of developmental system. I will use one of Griffiths and Gray's examples of this last sort of thing, involving two beetle morph that I call macho morphs and mini morphs:

> The successful developmental systems in certain beetle lineages have been those which produce one outcome in response to one sort of interaction, and another in response to a different interaction. The first produces a large, well-armed morph, the second a smaller morph that avoids conflict. Morphs of one type regularly give rise to the other morph. Both morphs are expressions of the same developmental system. (1994, 296-97):

So, clearly, when it comes to being a part of the developmental system, some polymorphisms are in and some polymorphisms are out. Beetle morphs, yes; deafened robins, no.

If a trait must be heritable to form part of an evolutionary developmental system, and both the mini and the macho morphs are parts of the same developmental system, there must be a way to conceive of the macho and mini morphs as expressions of a single heritable trait. This is not hard to do: the feature that is reliably replicated among the members of the beetle lineage is the trait of interacting with its environment so as to become either a mini morph or a macho morph. Griffiths and Gray are explicit that the developmental process can include discontinuous variation of this sort (1997, 476), and of course the beetle morph example is theirs.

However, the fact that we can conceive of heritable traits as encompassing discontinuous variation in this way is just the problem, for it will always be possible to take it that two distinct morphologies are versions of a heritable trait for which there is discontinuous variation. It will equally always be possible to take two distinct morphologies and treat the more common one as the heritable trait and the other as a developmental abnormality. What the notion of heritability does not tell us is when we should do which. Why treat the two morphologies of human wrist, scarred and unscathed in one way, while treating the two morphologies of the beetle in another way? After all, the two situations could be treated as parallel cases: sometimes humans in our lineage end up with scars on their wrists, and sometimes they do not; sometimes beetles turn out macho and sometimes mini, so we could treat the trait, wrist-scarred-or-not-wrist-scarred, as a heritable trait in the human developmental system, a trait expressing discontinuous variation. At least the restriction that the features of a developmental process must be heritable will not keep us from treating the human developmental system in this fashion.

Generally, if there are some occasions where one of two distinct interactions is considered part of the developmental system while the other is not, and there are other occasions where both of two distinct interactions fall inside the developmental system, the notion of inheritance cannot help us tell these situations apart. Such cases will have to be told apart; otherwise DST will collapse into the sort of holism that Sterelny found lurking in the DST program when he proposed his Elvis Presley problem. Everything that makes an individual developmental system distinct could be treated as an instance of a polymorphic trait for that sort of evolutionary developmental system. The result of treating all variations as versions of a heritable trait would be to construe the evolutionary developmental system as including all the developmental resources that produce any outcome exhibited by any individuals that are instances (or parts) of the system.

I should now note that Griffiths and Gray do indeed have a way of telling these cases apart. Recall that advocates of DST give two ways of delineating developmental systems. Developmental systems can successfully be delineated by appeal to adaptive historical explanations. The wrist-scarring interaction is not a part of the human developmental system because there is no adaptive-historical explanation for the wrist-scarred outcome that is Paul Griffiths (Griffiths and Gray 1997, 476). In contrast to the cases of the scarred wrist, the deafened robin, and the autistic human, both the two beetle morphologies, the macho-morph and the mini-morph, have adaptive-historical explanations, so that's why we should treat the beetle developmental system as encompassing this sort of discontinuous variation.

The problem with this second evolutionary criterion is that it makes DST's statement of the requirements for selection circular: in order to assess whether or not DST's putative requirements for selection are fulfilled, we must be able to delineate developmental systems

using our understanding of what does not and does have an adaptive-historical explanation, which presupposes an understanding of selection theory. This is essentially the same criticism as I deployed earlier with respect to Griffiths and Gray's criteria of individuation.

2.6 SUMMARY: THE PROBLEMS FOR RESEMBLANCE SELECTIONISM

In summary, the chief obstacles besetting resemblance selectionist approaches to the requirements for selection are the problems posed by 1) the notion of inheritance, 2) the delineation/individuation of appropriate macrosystems that must exhibit inheritance relationships, and 3) the grouping of entities into populations. My criticisms are these: The stances that the inheritance of fitness and the inheritance of phenotypes are requirement for selection are susceptible to counterexample; it is unclear how the macro-systems that are supposed to exhibit heritable traits are to be picked out; and, in order to make sense of the notion of inheritance, we have to have already grouped macro-systems into populations in the right way so as to use selection theory to make good inferences about system dynamics, meaning we already have to understand something about the sorts of systems over which we can deploy selection theory before understanding its requirements.

As we will see in the next chapter, the type selectionist approaches I consider below do not confront these difficulties. Inheritance is not used in stating the entrance rule for the theory, and how the entities that meet the entrance rule for the theory make up populations along with how macro-systems such as organisms can be delineated are presented later on, once the entrance rule for the theory has been specified.

3.0 TYPE SELECTIONISM

I now turn to consider Dawkins' type selectionist stance on the entrance rule for selection theory alongside my own type selectionist proposal. Recall that a stance on the entrance rule for selection theory is type selectionist if it requires that systems over which we deploy the theory feature entities that produce descendants of the same type as themselves. Dawkins' definition of the active germ-line replicator amounts to a stance on the entrance rule for (something much like) selection theory; my definition of the competitor is explicitly such. I consider my proposal alongside that of Dawkins because of what they have in common. Both are type selectionist, both avoid deploying the notions of inheritance and population, and both avoid reference to macro-systems.

The ways replicators and competitors are different have to do with features of Dawkins' definition of replicator that I think are either superfluous or implicit in his broader discussion of replicators. Dawkins both deploys criteria in his entrance rule for selection theory that I deem unnecessary and leaves implicit certain features of replicators that should be brought out into the open. Despite my difficulties with some aspects of Dawkins' view, and despite the superficial dissimilarities between the definitions of replicator and competitor, it should be acknowledged that the definitions that Dawkins and I propose will capture nearly all the same entities. The differences in our definitions of the entrance rule for selection theory should not obscure a broad agreement on how the deployment of selection theory is provoked.

In the next section, I set out definitions of both the replicator and the competitor. I then quickly discuss some objections to the notion of competitor that might immediately spring to mind in those who encounter the definition. From there, I consider Dawkins' account of the entrance rule for selection theory more fully, and use that discussion to motivate my own take on the entrance rule for selection theory.

3.1 REPLICATORS AND COMPETITORS

Replicators are anything in the universe of which copies are made. Germ-line replicators have the potential to have indefinitely many descendants. Active replicators have some causal influence on their probability of being copied (Dawkins 1982, 83). According to Dawkins, natural selection will occur whenever we find active germ-line replicators. Translated into the idiom of this work, Dawkins' claim is that selection theory may be deployed upon systems of active germ-line replicators. Note how the view is type selectionist: replicators produce descendants of the same type as themselves, not descendants that resemble them.

My notion of competitor is meant to function as an entrance rule for selection theory too, but I define it differently. An entity is a competitor by virtue of what it does: competitors struggle for existence with other competitors. We pick out competitors by finding entities that are struggling with existence with one another; entities that do so I call *rivals*. In order to pick out competitors in these terms, we need a precise notion of what relationship two entities have when they struggle for existence. Building on Darwin's discussion of the struggle for existence in the *Origin*, I claim that two entities struggle for existence with one another if each performs activities that both cause their own descendant production and inhibit (negatively cause) the descendant production of the other entity.

A paradigm example of the sort of activity that institutes the struggle for existence the consumption or monopolization of a limited resource, such as food. Darwin's first example of the struggle for existence in chapter 3 of the *Origin*, which is dedicated to consideration of that notion and how it functions as a prerequisite for selection, involves two dogs competing for food in a time of dearth (Darwin 1988[1859], 62). If there is only so much food to go around, and one individual in a population consumes some of it, this will not only facilitate the descendant production of the consumer, but it will also make it more difficult for other population members to produce descendants, since there will be less food for them.

More generally, competition of the sort I define above may arise in a number of other ways. Avoidance of a predator leaves it circling and hungry, and therefore raises the probability that other prey organisms nearby will get eaten instead. Taking a partner as a mate decreases the number available to others, too. These, too, are examples discussed by in the third chapter of Darwin's *Origin* of the sorts of relationships that institute the struggle for existence (1988[1859]).

If we pick out entities using the relationship of competition, we get a rather awkward definition. To make the definition more tractable, I will mean by "descendant production" to cause the production of descendants that are of the same type as the individual that produces them. Equally, I will mean by "inhibit" negatively cause. I also make heavy use of very general vocabulary, specifically "entity" and "type," and I promise to discuss this usage below. Here is the definition of competitor:

Competitor – An entity that i) produces descendants ii) does by means of at least one activity that inhibits the descendant production of a different type of entity, which iii) in turn produces descendants by an activity that inhibits the descendant production of the other entity.

While the definition may seem bizarre, the core idea behind the definition of the competitor is that competitors must struggle for existence to count as such. The activities that constitute competition are easily pictured using a directed causal graph:

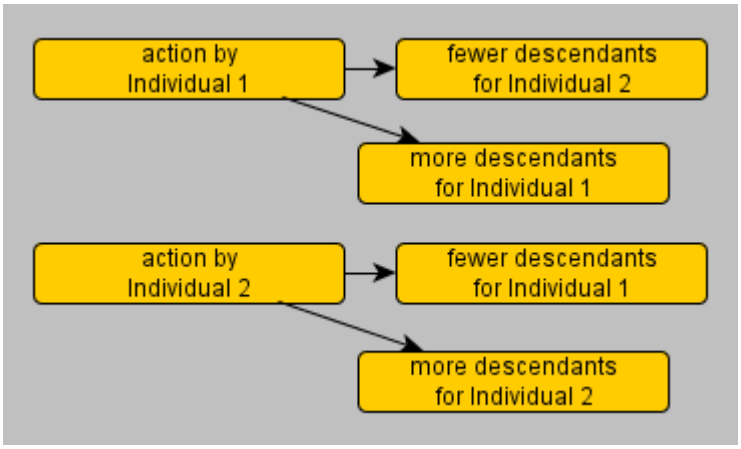


Figure 3.1 Graphical representation of competition

If any two individuals act in the above manner, both count as competitors. Below is an alternative presentation of competition, one in which the edges are weighted with a "+" for a positive causal connection and a "-" for an inhibitory connection:

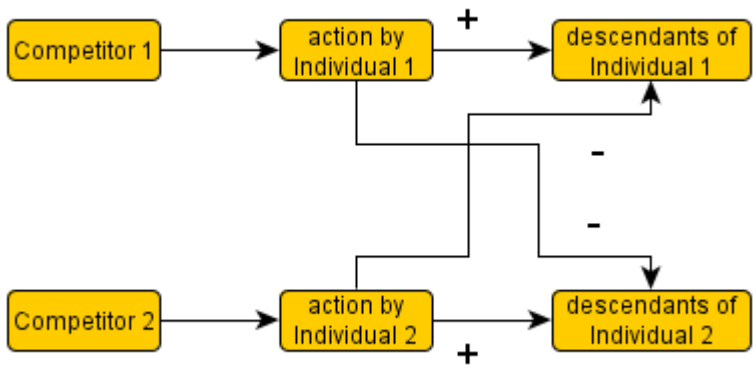


Figure 3.2 Alternative graphical representation of competition

3.1.1 Competition and the struggle for existence

The above characterization of competition is my own, but it is supposed to capture in causal terms the relationship that Darwin understood as the struggle for existence. Darwin himself does not define the struggle for existence; he gives the reader instances of the sort of checks on organic beings, and interrelationships among them, that create the struggle for existence. He thereby gives the reader an idea, indeed a very good idea, of what he means by the term.

Darwin discusses two cases of relationships that do not count as instances of the struggle for existence, despite being superficially similar to real cases of the struggle for existence, that I want to note. First, Darwin writes that a lone cactus at the edge of the desert may be said to struggle, but *not* in the sense in which creatures struggle for existence; "more properly it should be said to be dependent upon moisture" (Darwin 1988[1859], 62). I note this for two reasons. First, it highlights how struggling is a relationship between individuals. Second, Lewontin quotes this same passage as defense for the claim that the struggle for existence is not presented by Darwin as a necessary condition for selection (Lewontin 1970, 1). Lewontin misreads Darwin in doing so.

Darwin also writes that mistletoe, a parasite, does not struggle with apple trees, its prey, though several mistletoe seedlings on the same branch of an apple tree struggle with each other (1988[1859], 63). The dynamics of predator and prey populations do not in general resemble those of variant types of organisms, or variant alleles, as these are modeled in population genetics, so they had better be left outside the purview of the theory. We will make bad inferences about the dynamics of foxes and hares if we treat them as different types of individuals in a population genetics model, something we know because we know how to effectively model them using a different theory, evolutionary ecology.

Predators and preys are often said to compete; indeed, competition between individuals in the same species is often discussed alongside competition between predators and preys (e.g., Leigh 1999, 15). However, predators and preys do not compete in the sense specified above. Preys do not inhibit the reproduction of predators; rather, they promote their reproduction by serving as food for them. One might suspect that competition exists between predators and preys nonetheless because preys often deploy defense mechanisms to avoid their consumption. Are not predators and preys competing when such defense mechanisms are deployed?

In order to see that the answer to that last question is "no," we must divide up the cases of the deployment of defense mechanisms, say, flight from a fox by a hare, into two kinds: cases in which the hare flees the fox but the hare is caught anyway, and cases in which the fox merely chases the hare, but does not catch it. In the first sort of case, the hare does not impede the descendant production of the fox, because the hare serves as the fox's food. In the second sort of case, the fox does not cause its own descendant production by merely chasing the hare.

Contrast the relationship between two mistletoe plants parasitizing the same tree. These inhibit one another's reproduction as they each draw on the same limited resources provided by the tree. Each mistletoe draws resources from the tree thereby not only benefiting itself by acquiring the resources, but also inhibiting the reproduction of its rival.

Finally, note that population members compete even in very one-sided contests. The competition criterion does not require that rival types in a population fight one another to a standstill. If a heavily favored type spreads throughout a niche a little more slowly than it would have had it had no rivals, the process by which the favored type displaces the disfavored one may still be counted as an instance of selection in a single population. The loser in the competition nonetheless inhibited the descendant production of its rival, however weakly, even though its

causal influence over the process was overwhelmed by that of its rival. Generally, a process may terminate in a single outcome despite the impact of multiple causal influences. One should not confuse component causes with the results of causally complex processes to which they make a contribution.

3.1.2 "Type" and "entity"

I should briefly discuss my use of the notions of "type" and "entity" in the above definition. I am not the only person to use these terms in a specification of something like an entrance rule for selection theory. Dawkins defines replicators in terms of copying, and Szathmáry writes that entities qualify as units of evolution if they meet these criteria:

1. Multiplication. Entities should give rise to more entities of the same kind.
2. Heredity. Like begets like; A-type entities produce A-type entities; B-type entities produce B-type entities, etc.
3. Variability. Heredity is not exact; occasionally A type objects give rise to A′ type objects (it may be that A′ = B) (Szathmáry 1999, 31)

The concepts of entity and type may strike the reader as dangerously unspecified: How are we to decide what constitutes a different type of entity? In the case of alleles, our paradigm competitors, a rival type will have a different chemical structure. But the definition is supposed to apply more broadly. Griesemer raises this issue with respect to Maynard Smith's stance that evolution requires multiplication: "The precise nature of this condition on multiplication is vague. Any two things are similar in an indefinite number of ways, so every two objects produced by multiplication are of the same kind" (Griesemer 2000, 73). The threat here is that selection theory might be triggered too easily, most threateningly by populations of entities whose dynamics turn out not to be what selection theory would lead us to expect. Systems undergoing the wrong sorts of dynamics might be taken to be made up of competitors, if

"competitor" is defined in such a way that it is too easy to say what makes ancestors and descendants instances of the same type of entity. So how are "type" and "entity" to be used?

On the approach to that question pursued here, variant types of entities must compete, that is, engage in a specific sort of causal relationship, for it to be appropriate to deploy selection theory to explain their dynamics, but I want to leave open what other features the variants might have. That is the reason I use the terms "type" and "entity" in the above definition; I do not want use anything but causal features of systems to constrain the application of the term "competitor." It would a mistake to replace the notion of "type" in the above definition with "type of …" for fear that whatever phrase was used in place of the ellipses would circumscribe the application of selection theory too narrowly.

Officially, then, one can distinguish competing entities on any grounds, one can classify them into different types in any way, provided that the causal relationships specified in the definition hold of them. It is really the causal criteria that function to constrain the application of the definition; "type" is meant to be accommodating, to allow the theory to be deployed over variations that are not genetic variations or indeed any other more determinate type of variation.

Indeed, "type" functions in the above definition much as does the word "thing" in other contexts. If we want to talk about everything that shares some feature that is picked out using an adjective, such as "red," we often have to couple "red" with the word "thing"; for instance, if I want to say that no matter what other features something has, if it is red then it makes you hungry, I write "red things make you hungry." It's not that in order to make you hungry something has to be red and something else too, both red *and* a thing. Rather, "thing" is playing the role of a placeholder, allowing me to make the desired assertion about the relationship between redness and hunger without further limiting the sorts of things that make one hungry

beyond the constraint that they be red. When it comes to properties, the analogous notion for "thing" is "type" (and maybe "sort"); "thing" and "type" are both maximally general notions. Just as yellow garages are yellow buildings are yellow structures are yellow man-made objects are … are yellow *things*, red lenses are colored lenses are tinted lenses are visible lenses are … are *types* of lens. I want the causal criteria to do all the work in picking out competitors, so I forced to use words like "entity" and "type" in the definition to achieve the sought after level of generality.

Accordingly, those who would assail my entrance rule are invited to use, indeed even to abuse, the notion of type as they see fit in attempts to develop counterexamples to my position. My expectation is that they will find that while it is easy to come up with silly ways of distinguishing between different types of entities, but that not just any way of carving entities into types will yield ones that have the sorts of causal influences that the definition mandates. In particular, types that produces descendants of the same type as themselves will be hard to find, and ones that compete still harder. The causal criteria will restrict the application of the definition to thwart the proposed counterexamples, and that is how things are supposed to go.

3.2 REPLICATORS AS FULFILLING THE REQUIREMENTS FOR SELECTION

I turn now to consider more closely Dawkins' definition of the replicator. One must trace out some inferential connections in order to show that by saying that the active germ-line replicator is the "unit of selection," Dawkins can be interpreted as making a statement about the requirements for selection, or as stating an entrance rule for selection theory. I make this case first before I go on to discuss how well Dawkins' notion of "replicator" functions as an entrance rule for selection theory.

First and foremost, Dawkins' definition of the active germ-line replicator is meant to pick out the beneficiaries of adaptation. Replicators are *optimons* (Dawkins 1982, 84), and optimons are in turn characterized as "the 'something' to which we refer when we speak of a adaptation being 'for the good of' something" (Dawkins 1982, 81). Dawkins thinks of optimons as crucial contributors to teleonomy, what G. C. Williams (1966) proposed as the science of adaptation. Dawkins writes: "The central theoretical problem of teleonomy will be that of the nature of the entity for whose benefit adaptations may be said to exist" (Dawkins 1982, 81).

Wherever one has adaptations one should expect to find their beneficiaries, what Dawkins eventually picks out as active germ-line replicators. Adaptations must nearly always be the result of evolution by natural selection, and it is by virtue of this connection that the link between selection theory and Dawkins' specification of the *optimon* is secured: "The reason active germ-line replicators are important units is that, wherever in the universe they may be found, they are likely to become the basis for natural selection and hence evolution" (Dawkins 1982, 84). Dawkins' definition of the active germ-line replicator (henceforth just "replicator"), can accordingly be understood as a statement of the requirements for selection: It is where we find replicators that we find selection and adaptive evolution. It is in this sense that replicators are supposed to be *fundamental* to selection: wherever you get selection, you have replicators. While we might find adaptations that can be understood as for bodily survival or even for group survival, "all these adaptations will exist, fundamentally, through differential replicator survival. The basic beneficiary of any adaptation is the active germ-line replicator, the optimon" (Dawkins 1982, 85)

Unlike the many of the writers considered in the previous chapter, Dawkins does not make the mistake of treating evolution as a requirement for selection because there is nothing in

the definition of replicator that implies differential replication. First and foremost, replicators are what are responsible for *selection*, though Dawkins' way of putting this is a little odd: he has replicators being the beneficiaries of adaptations. However, there are other senses of "beneficiary" according to which other entities might be counted as beneficiaries of adaptations (Lloyd 1992). For this reason, I simply treat replicators as the things that institute a selection regime, or better, the things that trigger the deployment of selection theory, and leave behind the slippery language of benefit and adaptation.

3.2.1 What replicators do *not* do

It is worth noting that once one has picked out the entities that trigger the application of selection theory, one has not picked out either of two other sorts of things. First, one has not picked out *all* the entities that must be picked out in order to say how selection theory should be modeled in a given population. Second, one has not picked out entities that replicate autonomously, that is, without depending in any way on their causal context. Dawkins' claim that replicators are the units of selection has been criticized because there are entities besides replicators that must be invoked in selectionist explanations, and because replicators cannot replicate without their causal context taking very specific forms.

That Dawkins has been criticized in these ways is at least partly his own fault because he claims that replicators are the units of selection, and the notion, "unit of selection," has an unstable meaning. Dawkins uses the term "unit of selection" to refer to those entities that are the beneficiaries of adaptations, while Lewontin uses it to refer to another sort of entity, populations of which may evolve by natural selection when certain conditions are met. Indeed, if anyone is entitled to say what "unit of selection" means, then Lewontin gets the benefit of the dubbed, since he was deploying the term "unit of selection" long before Dawkins was (Lewontin 1970; Franklin and Lewontin 1970).

Nonetheless, because Dawkins puts forth the replicator as the unit of selection, he has come under criticism for the failure of his definition to pick out entities that are regarded as units of selection when that term is used differently. Dawkins has been criticized for claiming that selection is at the level of the gene when it is genotypes that are assigned relative fitness coefficients (Sober and Lewontin 1982; Sober 1987). Similarly, Griffiths and Neumann-Held point to models in which multiple alleles contribute to phenotypic variations that have a more proximate impact on organismic survival and reproduction than do the genes that cause them in an effort to undermine Dawkins' view (Griffiths and Neumann-Held 1999, 661). Dawkins has also been accused of giving alleles a sort of mythical ability to replicate themselves without having to rely on the scaffolding provided by various bits of cellular and environmental context (Lewontin 1991, 48). As I interpret Dawkins, the aim of his definition of the replicator was neither to provide a description of what sorts of entities can or must be assigned relative fitness coefficients in selection-theoretic models, nor what sorts of entities or features have the most causally proximate relationship to descendant production, nor what sorts of things can produce descendants independently of their causal contexts.

With respect to the first criticism, it should be acknowledged that there may be more entities at work in the process of selection than just those that need to be found to trigger the applicability of the theory. Most important among these additional non-replicators are those that bear relative fitness coefficients, principally gametes, zygotes, and mating pairs. While gametes, zygotes, and mating pairs (or generalizations of these notions) *do* play a role in selection theory, it is not the same role as that played by replicators and competitors. Replicators/competitors are what you must have in order for selection to go on. One need not have gametes, zygotes, or mating pairs or indeed any sort of interactors that are *distinct* from replicators in order for selection to occur (Hull 1988, 409). Rather, one gets different *sorts* of selection processes in populations featuring these different sorts of entities.

While the sorts of entities that deserve relative fitness coefficients in selection theory can indeed be defined (see chapter 5), this definition is different from the one that is used to trigger the deployment of the theory. Indeed, this is just what one would expect, since one needs already to have a grip on the notion of allele in order to say how gametes, zygotes, and mating pairs are recognized and differentiated in classical population genetics, because these individuals are distinguished in terms of the alleles they bear. Indeed, the trio of notions, "gamete," "zygote," and "mating pair," at least as they are used in selection theory, are inferentially dependent upon the notion of allele or some generalization of it for this reason.

Provided it is recognized that there are different sorts of entities that play different roles in selection theory and which need to be picked out using different criteria, it does not matter what we call the units of selection, whether its alleles, zygotes, mating pairs, or generalizations of these notions. All parties in the units of selection debate were on to something, and they all picked out as units of selection entities that play key roles in the theory. Instead of saying that Dawkins' replicator is the unit of selection, we should say instead that "replicator" is a generalization of "allele" (so is "competitor").

Griffiths and Gray provide a nice version of the second, "mythical self-replicating genes" criticism that they think works against even the relatively permissive extended replicator selectionism of Sterelny et al (1996). Griffiths and Gray claim that replicator selectionism is flawed because it will have to take into account non-replicators in order to explain how replicators replicate:

> The extended replicator theory also has to model the standing features of the physical world which form part of most developmental systems. Sunlight, gravity, mineral concentrations in the local soil, and many other factors must be present if "channels" are to convey and "replicators" to replicate. (Griffiths and Gray 2001, 197)

In response, it should be acknowledged that by picking out the entities that trigger the deployment of selection theory, replicator selectionists have not described everything that needs to be in place for such entities to produce descendants. Alleles require a host of contextual resources in order to replicate themselves, including standing features of the world that are not alleles, replicators, competitors, or even organic in nature. But insofar as Dawkins has picked out replicators as *things that get copied*, he has effectively, as a matter of logic, picked out only those things for which the scaffolding necessary for their descendant production is in place. An allele shorn of the developmental context necessary for replication is not, *by definition*, a replicator.

Things are much the same with competitors: insofar as I pick out competitors as things with causal effects on their own descendant production, a competitor that is an allele must, again by definition, have the necessary contextual scaffolding in place to act as a causal influence on its descendant production; without the necessary causal scaffolding, a would-be competitor will have no causal influence on the production of entities of the same type and hence will not count as a competitor at all.[11] In sum, Dawkins is best interpreted as offering his definition of the replicator as a statement of what sorts of entities can be expected to trigger the deployment of selection theory, and the conditions for the application of the theory of natural selection are stated fairly well using his notion of the active germ-line replicator.

[11] This means that a potential competitor that is not one until conditions change, say, until is begins to struggle for existence when a rival invades its niche, does not trigger the deployment of selection theory.

3.2.2 Biological vocabulary

Both Dawkins' proposed entrance rule for selection theory and mine eschew biological vocabulary, leaving it up to applicants of the theory to decide whether some biological structure should count as competitor or a replicator. This procedure can be justified in a couple of ways. For one thing, we should not "prejudge the empirical issues" and decide ahead of time whether some biological structure is fit for playing the role of the replicator or the competitor in selection theory. Furthermore, nearly every biological structure is the product of evolution by natural selection, including genes (Maynard Smith and Szathmáry 1995); insofar as we want a theory that makes such structures explicable, we should not presume that structures of these specific biological sorts had to exist for selection to go on (Okasha 2006, 15). A statement of an entrance rule for selection theory that deploys no biological vocabulary at all is structured so as to have at least the *potential* to be general enough to apply to any sort of biological structure, including both primitive ones that were once essential to the evolution as life on Earth but now may no longer exist, as well as novel ones that may yet arise.

While I eschew biological vocabulary in this work generally, one specific bit of biological vocabulary deserves special mention, namely "gene." Beyond the benefits of generality, I avoid this term for a couple of additional reasons. One is that, even in the context of modern population genetics, "gene" is systematically ambiguous, sometimes meaning "allele," and sometimes meaning "genetic locus." The other reason is that, while alleles are paradigm replicators/competitors, "genetic locus" is a much more difficult notion to understand.

Whether genetic loci can be individuated, and if so, how, is a much-debated topic in both the philosophy of biology and molecular genetics. I take no stand in this debate. I do, however, claim that it is easy to understand "allele," despite the fact that "gene" (in the sense of genetic

locus) is hard to specify. “Allele” can be understood more easily than “gene” because we need not pick out alleles by first picking out a genetic locus, and then subsequently picking out alleles as variations in the DNA at that locus. Rather, we can license our talk of differing alleles by first picking out genetic variations and circumscribing them afterwards.

Once genetic variation has been found, an individual who would apply selection theory to that variation must draw borders around it somehow. But we can leave decisions about how to draw those sorts of borders to those applying the theory, thereby allowing local considerations, including pragmatic ones, to dictate how to carve alleles from chromosomes exhibiting genetic variations. For instance, some approaches to multilocus modeling are simplest when there are only two alleles at each locus, a set-up that can be most easily achieved when genetic loci are treated as single nucleotide sites (Kirkpatrick, Johnson, and Barton 2002, 1731).

Furthermore, there is no reason to limit what sorts of information can be used to circumscribe alleles, either, and especially no reason to ban the use of information concerning *where* the variations are located on a chromosome. Griffiths and Neumann-Held argue that because selection may be prompted by variation at a single nucleotide site, replicator selectionists such as Dawkins are committed to thinking of single nucleotides as subject to selection, something that Griffiths and Neumann-Held think damns the position by *reductio* (Griffiths and Neumann-Held 1999, 660). But recognizing that sickle-cell anemia results from variant nucleotides *at a specific chromosomal location* where the nucleotides have definite distinct influences on development need not force one to acknowledge that variant nucleotides, when considered absent such positional information, form rival replicators or rival competitors.

I do not know whether single nucleotides count as replicators in Dawkins’ sense of replicator, since I am not sure how to evaluate the question of whether or not they are “active”;

but absent positional information that restricts them to a specific chromosomal site, single nucleotides do not count as competitors since they do not have inhibitory causal effects on one another's descendant production: cytosine does not, in general, have an inhibitory causal influence over the replication of guanine. And even if it did, that would mean a researcher would simply be *licensed* to take up the rather arduous task of deploying a selection-theoretic model to make inferences about the dynamics of these two competitors. Only a demonstration that such a model would imply a dynamics that are inconsistent with the real dynamics of the nucleotides, as assessed on independent grounds, would constitute a *reductio* of the position endorsed here.

3.2.3 The point of the entrance rule definition

Before we get to my proposed amendments to Dawkins' view, I should note one last point of agreement. One should judge the correctness of the definition of an entrance rule for selection theory by reference to the implications that such entrance rules are supposed to have. The goodness of any definition is a matter of how well it fulfills the *point* of the definition. In invoking the various criteria for active germ-line replicators that Dawkins does, he has his eye on the *point* of his notion of replicator. He writes in one place that "the whole purpose of our search for a 'unit of selection' is to discover a suitable actor to play the leading role in our metaphors of purpose" (1982, 91). In a similar vein, Dawkins remarks that it does not matter whether genes can be circumscribed according to some unitary criterion, for instance, one that would equate genes with cistrons: "My unit of selection, whether I called it gene or replicator never had any pretensions to unitariness anyway. *For the purposes for which it was defined*, unitariness is not an important consideration" (Dawkins 1982, 86; my emphasis). In my criticisms of Dawkins that follow, I argue on the basis of the point of the definition of an entrance rule for selection theory in order to establish that the competitor is a better concept to use as an entrance rule for selection theory than is the replicator, and Dawkins and I agree, it would seem, on how such arguments should work.

3.3 DAWKINS' DEFINITION OF THE REPLICATOR

I now turn to consider Dawkins' definition of the active germ-line replicator as a candidate entrance rule for selection theory. I find fault with every aspect of his definition, but the difficulties I find are small and easily repaired. Mostly, Dawkins' definition is too strong: some of the conditions he places on the entrance rule for selection theory are unnecessary and some of

them could be better stated using different language that allows the more general deployment of the entrance rule, generality that Dawkins almost certainly intended his concept of replicator to have.

3.3.1 The germ-line and potential immortality

In this section, I argue that the germ-line requirement in Dawkins' definition of the replicator is unnecessary. I argue that Dawkins implicitly recognizes this, too, as evinced by his discussion of individual replicators that are single nucleotides. Dropping the germ-line condition on replicators is harmless, and perhaps even beneficial.

As Dawkins defines them, germ-line replicators are potentially immortal, or at least have the potential to have very many descendants. (I will take it that "germ-line" and "potentially immortal" are synonymous, even though the former term has a narrower compass, being limited in application to genetic inheritance systems.) Dawkins' concern with the point of the theory of selection motivates his requirement that replicators be germ-line: somatic, "dead-end," replicators will not produce adaptations.

The criterion of potential immortality functions not only to exclude somatic replicators from triggering natural selection, but also to exclude replicators that are so large that they are bound to be destroyed by chromosomal recombination among polyploids. A very large chunk of chromosome that is bound to undergo recombination will fail to produce anything close to an indefinitely long line of descendants. Dawkins appeals to the point of Darwinian theory in discussing the problem of the size of replicators; our interest in deploying a theory that will make sense of adaptations is what functions to will limit the size of what Dawkins is willing to consider a replicator:

> So, *how* large and how small a portion of chromosome is it useful to treat as a replicator? This depends on the answer to another

> question: 'useful for what?' The reason a replicator is interesting to Darwinians is that it is potentially immortal, or at least very long-lived in the form of copies. (Dawkins 1982, 87)

So the requirement that replicators be potentially very long-lived in copy form is part of the entrance rule for Darwinian theory because it serves to narrow the focus of the theory to the sorts of entities that can produce adaptations. For Dawkins, it is a mistake to think of replicators in somatic cells as subject to selection because their lineages will inevitably terminate as the organisms that bear them die, and adaptation will not result from this sort of process; it is equally a mistake to think of massive chunks of chromosomes as replicators because they will produce very few descendants because of recombination. Such short-lived lineages do not contribute to the Darwinian science of making adaptation explicable.

The issue of the point of using the notion of replicator to make adaptation explicable comes up again in Dawkins' discussion of a possible *reductio* against his position, one based on the accusation that his definition makes it possible to treat chunks of chromosome that are *too small* as replicators. Specifically, Dawkins considers the possibility that single nucleotides fit his definitions of replicator (1982, 90). This is not a helpful way to deploy the term replicator, Dawkins responds. Crucially, however, Dawkins' criterion of potential immortality cannot help him forbid the deployment of selection theory over single nucleotides; the criterion he uses to ban chunks of DNA that were *too large* cannot help him forbid the deployment of the theory over entities that are *too small*. Instead, Dawkins relies on the pointlessness of the enterprise of using selection theory to explain the spread of single nucleotides *tout court*:

> The whole purpose of our search for a 'unit of selection' is to discover a suitable actor to play the leading role in our metaphors of purpose … I am suggesting here that, since we must speak of adaptations as being for the good of something, the correct something is the active, germ-line replicators And while it may not be strictly wrong to say that an adaptation is for the good of the

> nucleotide, i.e. the smallest replicator responsible for the phenotypic differences concerned in evolutionary change, it is not helpful to do so. (Dawkins 1982, 91)

This last response of Dawkins' casts into question the point of requiring that replicators be germ-line in the first place. While the criterion of potential immortality can rule out deployments of the theory over chunks of chromosome that are too large, ones that will inevitably be broken up in short order by recombination, it cannot rule out every non-serious, unhelpful deployment of the theory. Specifically, it cannot rule out the unhelpful deployment of the theory over replicators that are too small, such as single nucleotides. Instead, Dawkins must make a direct appeal to the point of his definition of the active germ-line replicator, rather than the concepts that constitute the definition, in order to rule against the selfish nucleotide. If such direct appeals to the point are sensible, then why not use them elsewhere too? Why not rule out the deployment of the theory over chunks of chromosome that are too big on the same grounds, that is, by directly appealing to the pointlessness of doing so? Were we willing to do that, we could drop Dawkins' germ-line requirement altogether.

Dawkins' instincts in responding to the "too little" problem are correct; his requirement that replicators be germ-line, which rules out replicators that are too big, is unnecessary. Generally, a good entrance rule for a theory need not rule out pointless applications of a theory, and that's what the germ-line condition does, at least in the cases where it is effective. Pointless applications of a theory may be a sort of mistake, but such mistakes cast the applicant of the theory in poor light, not the theorist. Unhelpful and pointless applications of a theory lead to foregone conclusions, not outright errors.

The measure of an entrance rule for selection theory is not whether or not it successfully bans pointless applications. Rather, a good entrance rule will not allow a theory to be applied

over systems that do not behave in the way that we infer they will when we deploy the theory (though we should not state the entrance rule in this way). When we apply the theory in a pointless way, we do not infer that systems will behave in ways that they actually do not; rather, we infer that they will behave in ways that we already know they will behave. That's what makes the application of the theory pointless.

This is easiest to see if one imagines actually applying selection theory to a chunk of chromosome that is too big, or to an allele in a somatic cell. The fitness of an "allele" that will inevitably be broken up by recombination is zero and hence its future relative frequency is too. So if we actually use the theory to infer the future relative frequency of an allele that is so large that it is doomed, we get exactly the result that we already knew to expect. Similarly, the equilibrium relative frequency of the descendants of variant alleles in different somatic cell lines is zero, too. These are foregone conclusions; I make no attempt to convince you of them because they are obvious. But they are exactly the conclusions one gets if one bothers to follow the rules of deploying selection theory in these sorts of easy cases.[12]

The individual who states an entrance rule for selection theory must state a rule that, if followed, will not invite someone to draw conclusions that are inconsistent with facts that are known on other grounds. Pointless applications of selection theory do not lead one to infer falsities of this sort. They need not be ruled out by a good entrance rule.

I discuss all this by way of argument for dropping the condition on replicators that they be potentially immortal. There are other good reasons for doing this too. "Potentially immortal"

[12] Strictly speaking, it might not be obvious how to reach the second conclusion by means of the machinery of selection theory. To infer the right long-term relative frequency of zero in the case of the somatic allele, one would require a temporally variable selection model that assigned the somatic cell allele an absolute fitness value of zero in the *n*th generation, where *n* is a reasonable guess at the number of potential cell divisions in the somatic cell line before death. Indeed, an entrance rule that allows one to fruitfully monitor the relative frequency of a somatic variation up to the *n*th generation but no further is not flawed for that reason. It may well be possible to use selection theory to explain the spread of rival somatic lines within the lifespan of an organism. Okasha, for instance, claims that we can think of cancers as outcompeting other non-mutant somatic cell lineages (2006, 11).

is a difficult notion to use in a rigorous and precise fashion. What's more, it may sometimes be valuable to deploy selection theory upon some system even if it is known that the rivals singled out will inevitably perish. Indeed, the fatalists among us will regard the lineages of all living things as doomed. Not only is the germ-line condition unnecessary, it might even be problematic.

3.3.2 Activity

Dawkins' replicator and my competitor are both defined through their causal relations. Dawkins' activity criterion requires that replicators exert phenotypic effects. I require that replicators exert a causal influence on their descendant production and on that of their rivals; I require that competitors compete. In this section, I argue that my notion of competition is better than Dawkins' notion of "activity" for the purposes of stating an entrance rule for selection theory.

Dawkins stipulates that replicators must be active, they must exert some influence over their probability of being copied (1982, 83). The use of the term "influence" should be interpreted as *causal* influence. This is clear from discussions in which Dawkins talks about replicators as exerting "phenotypic power" to increase in relative frequency (1982, 91), and his construal of cause-talk as difference-making (1982, 21). Including activity in the definition of replicators restricts the application of selection theory in such a way that neutral variations are excluded from fitting the requirements of the theory. But more interesting is how Dawkins' deployment of the activity criterion brings *rival* replicators into play.

Dawkins writes that his replicators are defined by reference to their alleles (1982, 92). Because replicators must be active, they must have causal effects, and these causal influences must be judged by reference to those of other replicators that vie for the same locus, hence alleles must have rivals. For Dawkins, it would be wrong to say that an allele at a locus at which there was no variation was active, because there would be nothing against which to measure the

allele's activity. What Dawkins' activity criterion involves, then, is the implicit presumption of *competition* between alleles. Indeed, Dawkins even refers to bits of DNA as competitors at one point (Dawkins 1982, 92).

My definition brings out the competitive nature of entities that trigger the deployment of the theory explicitly. This has the consequence that I do not rely on the individual applying the theory to make good decisions about what to count as a variant allele against which to measure the causal effects of a candidate replicator. Presumably, a replicator is not active if halfway across the world there is genetic variation at the same locus among members of the same species. It is clear, anyway, that the inferences we make about population dynamics using selection theory would not hold of a population in which the variants were separated from each other by an ocean. But there is nothing in Dawkins' entrance rule to *explicitly* eliminate the potential for this sort of misunderstanding. By requiring that competitors have rivals upon whose descendant production they exert inhibitory causal effects, I eliminate the possibility for this sort of misunderstanding and say explicitly against what sort of entity the causal effects of a competitor should be evaluated. My mutual inhibition of reproduction criterion plays much the same role as does Dawkins' activity criterion, but it does, I think, a better job of zeroing on the right sorts of systems over which to deploy selection theory.

My explicit reliance on the language of causation in the formulation of the definition of competitor has other benefits. Dawkins relies on cause-talk implicitly, since one evaluates whether or not a replicator is active on the basis of how its causal effects diverge from that of its rivals. However, Dawkins employs the language actual descendant production in his definition, "replicators are anything in the universe of which copies are made." He would have been better off with the language of causation here, saying that replicators are the sorts of things that cause

the production of their own descendants. Causal language is better here because effects need not happen for causes to cause them. Causes must influence effects, for instance by raising their probability. But since plenty of causes can be relevant to a single effect, it is not uncharacteristic of causes that they fail to actually produce their effects. Smoking need not produce lung cancer in everyone who smokes, yet smoking still causes lung cancer even among those who don't come down with the disease. Something *should* count as a replicator, and *does* count as a competitor, insofar as it is the sort of thing with a causal influence over its descendant production, whether or not it actually produces descendants. Dawkins clearly wants his notion of active germ-line replicator to be applied in this way: he writes that "a DNA molecule in the germ-line of an individual who happens to die young, or who otherwise fails to reproduce, should not be called a dead-end replicator" (1982, 83). But Dawkins must avail himself of the notion of potential immortality to secure this use of "replicator," while deploying cause-talk instead gets one the same conclusion without involving such rickety language.[13]

In summary, trading the language of "activity" for talk of inhibitory causal effects on rivals has only a few consequences for the entrance rule of selection theory: It legitimates the deployment of selection theory over neutral variations, and it makes explicit the contrast class against which the causal effects of candidate entities should be measured.

3.3.3 Copying fidelity

Another fashion in which my competitor differs from Dawkins' replicator is that I do not require any especially strong copying fidelity among the entities that trigger the deployment of selection theory. I require only that descendants be of the same type as their ancestors, and leave it to the applicant of the theory to find a means to distinguish the types. I am happy to deploy the theory

[13] While many might regard cause-talk as itself rickety, we are in business of using it anyway in explicating selection theory, so its use here does not add any indeterminacy.

over a population consisting of members of different non-interbreeding species such that the type differences are the species differences. Dawkins, on the other hand, will not regard an organism as a replicator, not even as a crude replicator with poor copying fidelity (1982, 99). Many organismic variations will not be passed on to descendants, so organisms cannot be replicators. Dawkins has been criticized for this view by those who think it at least possible that early life on Earth did not involve entities that had especially good copying fidelity (Maynard Smith and Szathmáry 1995; Griesemer 2000). The line of criticism is a good one: an entrance rule for selection theory should be sufficiently general that it can at least handle all episodes of adaptive evolution that have, or even *may have*, occurred.

I regard organisms as competitors, despite the fact that they exhibit poor copying fidelity. Provided it is possible to tell apart the descendants of the different types of competitors in the population, the lack of strong copying fidelity is no barrier to the deployment of selection theory upon the system. My position on this issue is much like that of Hull:

> Organisms behave in ways that make them candidates for replicators seldom enough without ruling them out by definition. As it turns out, in the most common situation in which one might want to view organisms as replicators— asexual reproduction via fission—it makes no difference. In asexual reproduction, usually the entire genome functions as a single replicator, and there is a one-to-one correlation between genomes and phenomes. Hence, the numbers will always turn out to be the same. (2001, 28-9)

The one-to-one correlation between phenomes and genomes in populations of individuals that do not interbreed is what makes it possible to be indifferent to whether or not genes or organisms are treated as replicators in the case that Hull considers: insofar as there is a sameness of type relationship between genomes there will be a corresponding one between organisms.

3.3.4 Gould's blending objection

S.J. Gould has questioned whether selection even requires entities that produce descendants of the same type as themselves. He notes that Darwin believed in blending inheritance, but also believed in natural selection:

> Blending inheritance marks an ultimate denial of faithful replication—for the hereditary basis of any selected character becomes degraded by half in breeding with an average individual. A paradox therefore arises. If units of selection must be faithful replications, and if Darwin both understood natural selection and believed in blending inheritance, then why did he ever imagine that selection could work as a mechanism? (Gould 2002, 622)

It would appear that by claiming that selection requires entities that produce others of the same type, I am ruling out the possibility of selection in cases in which organisms produce descendants that are a blend of the characters of their parents. That must be a mistake because Darwin, the first selection theorist, believed in blending inheritance. However, there is only an appearance of a mistake here, as becomes apparent when one considers that any form of such blending inheritance requires a mechanism.

We have already seen that the traits of ancestors do not have to covary with those of descendants for selection to go on within a population; more generally, the occurrence of selection is consistent with any sort of relationship between ancestor traits and descendant traits, including a blending relationship between them. So the case of "blending inheritance" should not pose special difficulties for selection theory. But any form of trait-level inheritance, blending or otherwise, requires a *mechanism* and it is properties of that mechanism to which one must look in order to adjudicate the question of whether selection occurs. Continuous variation of traits is compatible with discrete particulate inheritance of alleles. As Fisher (1918) established long ago, this makes sustained long-term selection possible despite the tendency of individuals with

extreme phenotypes to have less extreme descendants. Selection is possible despite a pattern of blending of phenotypes because the mechanism of inheritance is one in which ancestor alleles produce descendant alleles of the same type as themselves and do not blend.

I suppose it is possible to imagine systems in which there is blending of both phenotypes and hereditary material. However, it is not critical that an entrance rule for selection theory deem such systems subject to selection. When blending inheritance involves not only phenotypes but genotypes too, variation will rapidly disappear in a population (Fisher 1930, 5). A population that varied in this way is not one in which we could expect adaptations and other mysterious structure to persist; accordingly it is not one over which selection theory must be deployable so as to make adaptations or other mysterious phenomena explicable.

3.3.5 The blemish test

While Dawkins and I both endorse type selectionist approaches to stating (what I would call) an entrance rule for selection theory, unlike Dawkins, I allow organisms and other entities that do not make copies to prompt the deployment of selection theory. This is because I do not require that competitors pass what I call Dawkins' *blemish test*. Dawkins uses this test to justify singling out some developmental causes as replicators, paradigmatically alleles, and assigning these a special role in the study of natural selection (1982, 98-99). The test asks you to determine whether an entity whose structure has been modified will produce descendants that share the same modifications. Only replicators have this feature. For Dawkins, it is part of the meaning of "replicator" that replicators pass the blemish test: "The special status of genetic factors rather than non-genetic factors is deserved for one reason only: genetic factors replicate themselves, blemishes and all, but non-genetic factors do not" (Dawkins 1982, 99).

The connection between the blemish test and the requirements for selection is unclear. Why does it matter whether something passes the blemish test? The blemish test does rule out some things from counting as replicators, such as nests and other developmental products. Bateson, who first suggested that nests could be said to make nests by way of genes, recognizes that nests do not pass the blemish test while genes do (Bateson 2006). But why is it important to rule out nests? Perhaps variant nests *could* spread by natural selection, even if artificial manipulations of their structure were not replicated among their descendants. Provided variations in them are passed on from parents to offspring, there is the possibility of selection here, even if an individual with a mutilated nest variant produces offspring that construct a non-mutilated version. Whether or not *variations in the variations* get passed on seems to be a different question from whether or not variants produce descendants of the same type as themselves.

To be sure, structures that fail the blemish test are not the sorts of structures that we should in general expect to see bearing variations that spread by natural selection. Passing the blemish test is a requirement for a specific sort of inheritance mechanism to function as a recurring and vital source of variation on which selection can act. Because variations in nests do not generally reappear in descendant generations, we should not expect to see nests functioning as replicators, at least not very often. In contrast, DNA *is* the sort of thing we should often see functioning as a replicator because it passes the blemish test. So DNA will prove a recurring source of variation upon which selection can act while nests will not. But this distinction is orthogonal to the one between systems that meet the entrance rule for selection theory those that do not. Triggering repeated bouts of selection is not a requirement for triggering selection *tout court*.

Nothing is lost by refusing to use the blemish test to determine whether or not something counts as a competitor. The requirements already in place in the definition of the replicator will rule out variant nests from counting as replicators, for as a matter of fact they do not produce descendants of the same type as themselves. Besides, one can imagine intermediate mechanisms that merely perform so-so on the blemish test, ones that pass on some of their blemishes to descendants but not others, and there is no reason to refuse to countenance variations in these as fuel for selection.

One consequence of dropping the blemish test is that variant developmental products, as opposed to variant developmental causes, can function as competitors. Consider a simple scenario in which a variant haploid allele does nothing but produce a variant organismic trait. In this case, what are the competitors, the alleles, the traits, or the organisms with the variant traits? If we used the blemish test to answer that question, we would have the say that the alleles are the competitors, since one could manipulate the trait artificially and the manipulation would not be passed on to descendants. But if we refuse to use the blemish test, we cannot privilege the variant alleles as our rival competitors. The fact that we could manipulate the alleles such that the manipulations got passed on to descendants does not grant them special status. But when we deploy population genetics models for haploid populations, we get the same results whether we count variant traits, variant organisms, or alleles. This was just Hull's point in the quotation from the previous section. So it does not matter what we pick out as our variant competitors in cases of this sort.

While developmental products and organisms *can* function as competitors, there are more complicated cases in which we will be forced to pick out genetic variations as our competitors, rather than the phenotypic variations they produce. Among diploids, variant alleles produce

alleles of the same type as themselves, but variant organisms may have offspring of a different type than themselves. This occurs, for instance, in cases of absolute dominance when a recessive homozygote has heterozygote progeny. Classifying heterozygotes together with homozygotes as entities of the same type would lead one to make faulty inferences using population genetics models. With increased causal complexity, such as that produced by interactive environmental causes, it becomes increasingly hard to think of phenotypic traits as competitors in populations whose members interbreed. Descendants may mature in different environments than did their parents and express radically different traits, as in the counterexamples to the resemblance selectionist view offered earlier. So we will often be forced to treat the alleles as the competitors and not the traits when dealing with phenotypic variations in populations of diploids.

Sometimes, however, we can harmlessly consider organisms competitors, for instance when we apply selection theory to populations composed of members of different species, or, more generally, when we consider competition between individuals of different types that do not interbreed. Indeed, an early imaginary scenario that Darwin uses to help the reader understand the process of selection is one in which a local population must compete with new immigrants (1988[1859], 81).[14] Here is one case where Dawkins' blemish test will let us down. Organisms do not do very well on the blemish test, since their descendants often do not look exactly like their parents, and we will not always be able to use genetic differences between members of different species that are competing with one another as grounds for deploying selection theory to explain their dynamics. Darwin, anyway, was hardly in such a position. Still, we can know to apply selection theory to a system composed of members of different species in the fashion that Darwin imagined in the *Origin*.

[14] Darwin then modifies the scenario by having the reader imagine that the new forms are the result of local variation.

Competition between members of different species or clades may even require treating organisms as competitors. Indeed, the same mathematical models of population genetics can be used to explain the dynamics of populations of rival haploid alleles and rival members of different species, as is remarked in passing by de Meeûs and Goudet (2000, 982). Treating organisms as competitors in these sorts of cases is not problematic provided descendants and their parents can be classified as of the same type.

As in similar cases discussed already, dropping one of Dawkins' rules for picking out replicators requires the adoption of a new rule to perform the task that Dawkins' rule did. Dawkins had strict criteria for determining what it takes to produce descendants of the same kind, to be a replicator. Exact copying fidelity, the transmission of features "blemishes and all" is what it takes to produce descendants of the same kind. Descendants must be exactly the same as ancestors. If we do as I suggest and loosen this requirement, we face the question of how we are to classify competitors into types so that we can tell when selection will go on. As discussed earlier, I think we can be perfectly generic about how we classify competitors into types. The causal criteria in the definition of "competitor" should do the work of restricting the applicability of the definition; types can be distinguishes in any way.

In sum, then, Dawkins' replicators and my competitors are not all that different. Dawkins and I both think our definition can cover instances of "expanded inheritance" and can even be used to determine what sorts of expanded inheritance we should really expect to behave as we would infer they will when we deploy selection theory (Dawkins 2004). And even insofar as our definitions differ, they pick out mostly the same things, though I would claim my use of causal language is better than Dawkins' talk of potential immortality. As far as differences of

application go, my notion is can be deployed more broadly than Dawkins' because it could well be used on entities that do not copy themselves in an exact sense of "copy."

3.3.6 Inheritance and causally complex intermediaries

Several authors have offered criticisms of replicator selection, criticisms that, though directed at other writers (typically Hull and Dawkins) should be addressed here. Some of these attacks do not strike the competitor definition I offer because of how that definition is formulated differently from those of Hull and Dawkins. Some of the criticisms work only against those who make more ambitious use of the notion of replicator, those who would use the notion of replicator to do more than just trigger the deployment of selection theory. I have already addressed (section 3.2.1) two prominent concerns about type selectionist approaches to selection theory, specifically that these approaches overlook scaffolding causal influences over replication, and overlook the importance of entities that bear conglomerations of competitors. Here I consider some other criticisms of the replicator selectionist perspective.

Godfrey-Smith (2000) has claimed that replicators are not essential to inheritance and evolution on the grounds that the production of descendant copies of replicators may involve causally complex paths that are highly contingent. He imagines that reverse transcription is possible, such that DNA is never directly causally involved in the production of more DNA, but instead produces proteins from which DNA is later generated through reverse transcription. Indeed, he goes on to consider the possibility that there might exist systems in which the causal connection between ancestor replicators and descendant replicators is even more indirect than this. Reverse translation does not occur among Earthly organisms, but the replicator concept is supposed to be fully general, so the possibilities that Godfrey-Smith imagines are the sorts of things with which advocates of replicator selection must contend. I detect three separate difficulties raised by the possibility that Godfrey-Smith's considers.

First, Godfrey-Smith claims that the sorts of cases he imagines are ones in which there is a lack of a replication event in which an ancestor replicator produces a descendant replicate. That there is a complex causal process lying between a replicator and its descendants strains the notion of replication. Perhaps it does, but it does not matter whether or not ancestors produce descendants by way of a complex causal process or simple copying. Maybe "replication," "replicate," and "copy," were bad notions for Dawkins to use when devising his novel perspective on selection theory because in using them he would seem to be excluding attenuated descendant production processes, though I suspect he chose the notion of replicator because replicators pass the blemish test. Anyway, provided we are interested in using the concept of replicator for the purposes of saying under what conditions selection theory may be deployed, we need not formulate our entrance rule by appeal to simple copying, rather than a multistep attenuated process. Maybe causal chains of the sort Godfrey-Smith imagines do not constitute

ones in which *copies* are produced, but they are the sorts of things that can be trigger selection processes, provided they are *causal* chains (see Nanay 2002).

Another aspect of Godfrey-Smith's criticism is based on the notion of causal responsibility: when the causal connection between a replicator and its descendant becomes attenuated, and many other factors are found to be essential to descendant production, one can no longer isolate the original replicator as causally responsible for the production of its descendant:

> The more factors that are involved in creating a new Y that is similar to X, and the more places in the network at which dissimilarity could be introduced, the less true it is to say that "X was causally involved I the production of Y in a way *responsible* for the similarity of Y to X," as the definition requires. (Godfrey-Smith 2000, 19)

Godfrey Smith is right that plenty of other things will be causally involved in the production of descendants on the part of a single replicator or competitor, and that is enough to show that we cannot pick out the entities that trigger the deployment of selection theory as ones that have *something to do* with the production of descendant replicators. But replicators (implicitly) and competitors (explicitly) are singled out as special by virtue of being causally engaged in *competition*, and they can be isolated among the causally responsible factors for this reason. The scaffolding that makes descendant production possible is not engaged in the mutual inhibition of descendant production, or if it is, as in cases of multi-locus selection, then it is rightly counted as a competitor too.

Another nearby issue is creeping holism, the concern that a liberal notion of replicator will allow too many things to count as replicators, so that in the limit, there are no replicators but instead, "the entire causal network in the life cycle somehow manages to transmit variation from generation to generation" (Godfrey-Smith 2000, 19). Thumbs will count as replicators according to a liberal notion of replicator that requires only an attenuated causal connection between

ancestors and descendants (Godfrey-Smith 2000, 19). Godfrey-Smith writes that "as the causal web gets more complicated, it becomes less and less appropriate to try to identify a replicator, where a replicator is a definite entity, or identifiable lineage of related structures, that is responsible for heredity" (2000, 19).

Godfrey-Smith is right about heredity. Conceived as the recurrence of specific morphological or phenotypic features between ancestor and descendant entities, such as thumbs, inheritance is indeed explained by more than just replicators. In a similar vein, Jablonka and Lamb complain that "the replicator concept is associated with a very specialized type of information transmission, which does not cover all types of inheritance, and therefore cannot be the basis of all evolution" (Avital and Jablonka 2000, 359). They are right, too. Replicators can explain neither all evolution, nor all types of information transmission, nor all types of inheritance. The replicator and the competitor were never designed to explain recurrence of form (Dawkins 1982, 88); they were designed to pick out key entities in selection processes. Selection may be triggered by rival replicators/competitors even if inheritance, recurrence of form, transmission of information, and even evolution are not explained by them in some cases.

This same sort of response as the one just rehearsed can be leveled at some instances of expanded inheritance that have been offered as challenges to the replicator-selectionist view. For instance, Russell Gray has pointed to the track-and-bowl structures that kakapo parrots inherit from their ancestors in an effort to undermine gene selectionism. Gray claims that genic selectionists cannot adequately model how the males of these large flightless New Zealand parrots modify their environment so as to attract mates. The track-and-bowl adaptation is unusual in three ways: the birds actively modify their habitat; males that utilize the system are at a reproductive advantage; and the tracks and bowls are inherited across generations and recreated in novel environments (1992, 196).

This is surely an unusual form of inheritance, but it has nothing special to do with selection. Of course, that such structures got constructed in the first place probably has something to do with selection that occurred long ago: the parrots probably became disposed to modify their environments in the fashion that they do because some competitors, specifically ones that disposed their bearers to construct track-and-bowl structures, out-competed others in the ancestral lineages of the parrots. But that bout of selection explains track-and-bowl construction behavior; it does not feature tracks and bowls in the role of competitors.

Even though track-and-bowl structures are inherited these days and can vary such that they have an impact on the reproduction of the parrots that inherit them, those facts should not prompt one to deploy selection theory over variations in track-and-bowl structures. Generally, organisms may have especially many (or few) offspring because they happen to be systematically associated with sub-environments within a broader ecosystem. Mathematical models in population genetics have been developed to cope with these sorts of systematic influences (Christiansen 1975). This sort of systematic association between members of lineages

and ecological contexts is what is going on in the track-and-bowl scenario: better track-and-bowl systems are not *spreading* at the expense of others, rather they are *enduring* because they facilitate the reproduction of members of a lineage that use them and keep them in good shape. But what demystifies their occurrence and their perpetuation is the disposition of all parrots, both lucky parrots with good bowls and unlucky ones with bad bowls, to create and keep up track-and-bowl structures in general.

If you are in the business of demystifying the bowls, you need not, and indeed should not, feature them in the role of triggering the applicability of selection theory. One should not apply selection theory to the track-and-bowl systems as an instance of "expanded inheritance" in the way one might do with rival cell membranes. The bouts of selection that demystify the track-and-bowl structures have long passed. The real mystery is how these structures came to be in the first place, and that mystery is solved by showing how it could result from the spread of alleles that disposed their bearers to build them.

The type selectionist view should be understood as an alternative view to the resemblance selectionist one in which inheritance plays a key role in stating the requirements for selection. We have seen that using inheritance in this way is a mistake, so we should not be surprised to find that there are instances of inheritance that do not trigger the deployment of selection theory.

3.4 JUSTIFYING THE COMPETITOR

It is difficult to provide a full justification for my definition of the competitor as an entrance rule for selection theory. In proposing an explicit generalization of the circumstances of application of a scientific theory, I find myself in the standard situation of a theorist. Theorists cannot generally prove that their accounts will never invite anyone to infer false conclusions from true

premises. I am making a suggestion for how to circumscribe the domain of applicability of selection theory, but I cannot establish by a deductive argument that every system within that domain will be one over which it is appropriate to deploy the theory.

Still, the previous section, in which I discuss how the competitor compares Dawkins' replicator, does some of the work of making my proposed entrance rule for selection theory seem reasonable. And, unsurprisingly, the competitor fulfills the necessary conditions on an entrance rule for selection theory that I put forward in chapter 1. But the true measure of an entrance rule for a theory is whether it captures the inferential territory being sought. In the case at hand, that requirement amounts to this: it must be possible to show what an explanation of any adaptation, instance of altruism, or genetic polymorphism would look like by deploying selection theory as I formulate it, where the existence of rival competitors is an explanatory starting point for the deployment of the theory. Competitors trigger the application of the theory, and they had better trigger it in the right cases, including *all* the cases in which dynamics of the above sorts are possible.

While I cannot show that the definition is not subject to counterexample. I can, however, argue that what look like deficits of the proposal are not really such. For instance, it would appear *prima facie* that too many entities would end up rivals with one another if the competitor formulation of the entrance rule for selection theory is adopted. To mitigate this criticism, I draw attention to the special nature of the causal relationship that must exist between two rivals. The relationship is no ordinary causal relationship. Surely, lots of biological entities have causal effects on another, even effects on the descendant production of other entities. But to be counted as rivals, entities must have *mutually inhibitory* causal relationships on one another's descendant production, and they must promote their own descendant production as well. That sort of bidirectional inhibitory causal relationship is special, special enough, I hope, to bring together all and only those systems whose dynamics are explicable using selection theory.

It is perhaps worth noting that some initially plausible counterexamples do not turn out to be such. Consider that predators and prey do not count as rivals, as argued early on in this chapter. Consider also that alleles will compete only with different alleles that contend for the same place on the genome. Alleles at different loci will not compete. Even two deleterious alleles within a single organism will not compete, for though they each have deleterious effects on the descendant production of one another, they do not cause their own descendant production, but rather impede it. In short, the competitor definition is not as given to blowing up and including far too much as first appears.

On the other hand, the entrance rule that I suggest is fairly flexible and can include a wide range of inputs. For instance, while alleles are paradigm competitors, organisms that are members of different species can stand in the right sorts of competitive relations to trigger the deployment of selection theory too. The evolution of culture is the non-traditional arena in which

selection theory models, or at least models resembling those of selection theory, have been deployed most vigorously. Memes are hard to circumscribe, but it seems reasonable that at least some things that are supposed to count as memes will count as competitors too.

However, my entrance rule is not so flexible that it can accommodate every mathematical model that has been developed by researchers concerned with cultural evolution. The moon model of Henrich and Boyd (2002) will not fall under the purview of selection theory as I conceive it because it involves entities that produce descendants of a different type than themselves. I discuss that model at greater length in the conclusion.

Selection-like models of neural development in which neurons do not produce descendants, such as that advocated by Changeux and Dehaene (1989), also fail to count as selection theory models according to the entrance rule I offer because neurons do not produce entities of the same sort as themselves.[15] Selection-like models of the immune system would seem to fail to fall under the purview of selection theory, too, since there does not seem to be competition between differing antibody-producing B cells (Hull, Glenn, and Langman 2001). However, operant conditioning does seem to count as the sort of process that can be modeled using selection theory as it is construed here, since there do seem to be conditions in which different sorts of behavior mutually inhibit one another (Hull, Glenn, and Langman 2001). So some domains in which selection has been said to go on are ones that fit the entrance rule for selection theory as I conceive it, and some do not.

Ultimately, even if the entrance rule carves up the territory of such systems in a queer fashion, say, by including some cognitive development applications but excluding neurological ones, or including only *some* models of cultural evolution, the entrance rule need not be regarded

[15] A reader asks whether time-slices of neurons might produce time-slices of neurons, thereby qualifying models of neural selection as selection theory models. I am not sure whether this will work; I will look into the matter.

as totally inadequate for that reason. What is most important is that the entrance rule recovers at least the territory it was designed to capture, that is, all instances of selection that lead to adaptation, polymorphism, and altruism; whatever else falls under its purview is bonus material. Perhaps, too, someone else will come up with an even more general theory with an entrance rule that treats competitors as a special case but can also harbor immunology and other applications; I would welcome such a generalization.

However, not just anything will count as a generalization of selection theory. Specifically, the mere fact that the same mathematical models can be used to capture both changes in the relative frequency of competitors along with the dynamics of other sorts of systems is no reason to think that they all must be covered by the same theory. At least, this is the case for *syntactically* identical models. Earlier I argued that Lewontin's requirements for selection, which led to the exclusion of models of overdominance from the purview of selection theory, were flawed for that reason. Those models shared with all classical population genetics models variables that refer to the same sorts of things, relative frequency terms referring to different types of zygotes weighted by relative and average fitness parameters. The diffusion of heat in a wire is covered by the same diffusion equation that is used for the spread of alleles in large populations undergoing relatively weak selection, but that hardly implies that heat diffusion should fall under the purview of selection theory or that selection is an instance of thermodynamic change. Those models do not share anything more than syntactic structure. There is much more to deploying selection theory than just turning a mathematical crank. Only insofar as there are many significant commonalities between, say, the pruning process in neural development and the population dynamics of competitors would it seem odd to place these processes in different theories.

In the sections that follow, I discuss how competitors can interact in special ways so as to form gametes, zygotes, mating pairs, subgroups, and substructures. I discuss how they can be grouped into populations. I discuss how the causal influences on their dynamics can be botanized according to a set of criteria such that inferences can be made about their dynamics on the basis of that causal information. As far as I can tell, none of these features of deployment of selection theory apply in the case of neural pruning. With this in mind, the fact that the same mathematical model can be used to explain neural pruning as can be used to explain some instances of selection does not seem, all on its own, to provide very solid grounds for considering neural pruning an instance of a process that must be covered by selection theory.

As a final way of motivating my entrance rule for selection theory, I note the importance of competition in selection theory generally. Competition is important historically: Darwin made the struggle for existence a prerequisite for selection. But in the context of the specific approach taken here, it turns out that one can do much of the work of saying not only *when* but also *how* selection theory should be deployed using the notion of competition. Competition is used in what follows as part of the criterion for circumscribing populations in selection theory. It is used to justify the attribution of relative rather than absolute fitness values in population genetics models. It is used in the definition of the special sort of relationship that alleles in a zygote and mates in mating pairs bear to one another. It is used to contrast hard and soft variable selection models and to distinguish subgroups and substructures from other sorts of groupings. The picture that emerges in the story that follows is that one simply cannot say how to deploy selection theory without using the notion of competition to say it. Accordingly, competition needs to be a feature of any system that is picked out as one over which selection theory may be deployed. The easiest way to make sure that competition is a feature of those systems over which selection theory is to be deployed is to make competition part of the entrance rule for the theory.

4.0 "POPULATION" IN SELECTION THEORY

In the previous chapter, I offered a definition of what features a system must have in order for it to be appropriate to deploy selection theory to explain its dynamics. In this section, I begin my exploration of the "guts" of selection theory. The next four chapters aim to say how populations of competitors are circumscribed (this chapter); how causal features of those populations can be connected to deterministic Wright-Fisher population genetics models (the algorithm of chapters 5 and 6); and how those models can be transformed into stochastic models through the deployment of equations that take as inputs the outputs of the deterministic models coupled with a new parameter, *variance effective population size* (chapter 7). The order in which I present these issues is not haphazard: we must have already delineated the scope of a target population before we can apply the algorithm to it, and we must have already generated a deterministic Wright-Fisher model before we can deploy a stochastic version of one.

One of the arguments from a previous chapter turned on the claim that there are correct and incorrect was of grouping the entities to which selection theory applies into populations. I illustrated this by quoting Damuth, who considered an absurd putative population consisting of organisms living half a world away from each other. But if one can go about grouping entities into populations in the wrong way, one can go about doing so in the right way. The right way for grouping entities into populations is the way, or at least a way, of doing so such that were one to apply the rest of the apparatus of selection theory to that grouping, one would make correct inferences about its dynamics. Even if one follows all the other rules for deploying selection

theory correctly (determining gene frequencies, ascribing relative fitness coefficients, etc.), one will not derive from the theory good predictions or good explanations if one does not circumscribe one's population correctly. That is just to say that there is nothing about the rest of the rules for deploying selection theory that will keep selection theorists from deploying the theory over a poorly circumscribed population. We should not think of biological populations as a given either. The notion of population cannot be specified in some theory-neutral fashion (Gannett 2003).

Furthermore, as I argued earlier, knowing how to group entities into populations in the right way is a part of deploying selection theory correctly, just as ascribing fitness values in the right way is part of deploying selection theory correctly. The right way to group entities into populations for the purpose of selection theory is surely understood *implicitly* by selection theorists: no one would dream of forming populations whose members spanned different continents. An account that seeks to make selection theory *explicit* must include a *rule* for delineating populations.

As with the other rules proposed in this dissertation, I am constrained to state a rule for determining the scope of a population using causal vocabulary, along with statistical and logical vocabulary. Before proposing my own rule for circumscribing populations, however, I will consider some definitions of "population" offered by biologists and expose these as inadequate for our purpose. I will also discuss some insights that population geneticists and philosophers have made about populations in selection theory before presenting my rule for determining population size. From there I will draw out some of the subtler aspects of the definition and finally go on to defend it as suitable for the purposes of deploying selection theory.

4.1 POPULATION AND CENSUS POPULATION SIZE

My aim is not to offer an account of the notion of census population size (*Nc*), a variable whose value is determined on the basis of the number of breeding adults in the population. There are a few reasons for this. For one thing, census population size does not include the individuals who struggled for existence with the individuals who went on to become breeding adults, but failed to achieve reproductive success, and these individuals' struggle should be taken into account in a determination of why the population ended up with the number of breeding adults that it did. Counting only reproducing organisms, or the alleles within them, makes sense for determining a value for *Nc* as it functions in derivations of effective population size, a critical variable in stochastic dynamical models. But I am interested in picking out systems over which selection theory should be deployed, and this must include more than just the breeding adults. Furthermore, while breeding adults die out after reproducing, at least in populations exhibiting discrete generations, the systems over which selection theory is deployed must persist in time. So a criterion for the notion of population should exhibit populations as enduring sorts of things, whose members can come in and out of existence even as the population itself remains.

4.2 BIOLOGISTS' DEFINITIONS OF POPULATION

There has been a recent surge of interest among evolutionary biologists in the concept of population (Waples and Gaggiotti 2006, 1420). The work just cited is a review article of recent studies of population structure that have been spurred by the availability of polymorphic DNA markers in natural populations. The authors seek to evaluate different procedures for estimating the number of substructures that exist within a structured population on the basis of genetic samples from it. But before they embark on that project, Waples and Gaggiotti's discuss the notion of population in a general way, drawing on other biologists' characterizations of it.

Waples and Gaggiotti bring out several interesting features of biologists' use of the term, "population." For one thing, biologists interested in evolution have neither settled on a definition of the term "population" nor are they trying very hard to do so. Waples and Gaggiotti also note that, despite considerable heterogeneity in the proposed definitions, the problem of how populations should be circumscribed gets surprisingly little attention from biologists (2006, 1420). Though biologists do not fiercely defend any of the characterizations, Waples and Gaggiotti's work contains a helpful table of suggested definitions of the term "population" that have been proposed by biologists, one the authors claim is representative (2006, 1420). I'll review several of the definitions here, focusing especially on ones that fall under Waples and Gaggiotti's "evolutionary paradigm," but considering a couple from the "ecological paradigm" and "variations" too.

4.2.1 "Population" under the evolutionary paradigm

Here is a sample of the definitions from Waples and Gaggiotti (2006, 1420); these ones fall under what they call "the evolutionary paradigm":

- A community of individuals of a sexually reproducing species within which matings take place (Dobzhansky 1970).

- A major part of the environment in which selection takes place (Williams 1966).
- A group of interbreeding individuals that exist together in time and space (Hedrick 2000).
- A group of conspecific organisms that occupy a more or less well-defined geographical region and exhibit reproductive continuity from generation to generation (Futuyama 1998).
- A group of individuals of the same species living close enough together that any member of the group can potentially mate with any other member (Hartl and Clark 1988).

4.2.2 "Population" under the ecology paradigm

Here is a sample of the definitions said by Waples and Gaggiotti to fall under the ecological paradigm (2006, 1420); the last one is from the category "variations":

- A group of individuals of the same species that live together in an area of sufficient size that all requirements for reproduction, survival and migration can be met (Huffaker, Berryman, and Laing 1984).
- A group of organisms occupying a specific geographical area or biome (Lapedes 1978)
- A set of individuals that live in the same habitat patch and therefore interact with each other (Hanski and Gilpin 1996)
- Natural population: can only be bounded by natural ecological or genetic barriers (Andrewartha and Birch 1984)

4.2.3 Interactions in biologists' definitions of "population"

Waples and Gaggiotti write that all the definitions they survey, "imply a cohesive process that unites individuals within a population" (2006, 1421). They detect among workers with an interest in ecology a focus on the social, behavioral, and competitive interactions that are made possible by members residing in the same places and the same times; writers focused on evolution are more interested in reproductive interactions and shared genes (2006, 1421). Nonetheless, the definitions that fall under the evolutionary paradigm make reference to such things as proximity of population members, shared environments or geographic regions, and shared space/time locations. We can tentatively conclude, then, that biologists interested in population dynamics recognize the importance of interactions among population members and evolutionary biologists place special emphasis on reproductive interactions.

4.3 CRITICISMS OF BIOLOGISTS' DEFINITIONS OF "POPULATION"

The list of definitions of population drawn from Waples and Gaggiotti above is enough to show the importance of interbreeding as a criterion for delimiting populations in the minds of biologists. Shared ecological and time/space coordinates are also widely mentioned, something I call the geographic criterion. Some of the writers also appeal to taxonomic relationships to delineate populations. The criterion I offer for delimiting populations below is interaction-based too. However, the kind of interaction I postulate is different from those suggested by biologists and population geneticists.

In this section I criticize the ways biologists have suggested for delimiting populations in order to motivate my alternative. I note upfront, however, that I recover these criteria as convenient ways to operationalize the proposal I make for delimiting populations in selection theory, since my criterion cannot be deployed directly (see section 4.4.3).

4.3.1 Why Interbreeding criteria and geographic criteria will not work

The interbreeding criterion is of limited scope, applying only to creatures that propagate themselves sexually, so it cannot function as part of the general account of selection being proffered here. But the interbreeding criterion can be motivated in part by the fact that many models in population genetics explicitly make assumptions with respect to breeding relationships among population members. It is worth stressing, too, that interbreeding relationships are interactive relationships: to delimit populations using the interbreeding criterion is to delimit them in terms of whether their members interact with one another in a certain fashion.

4.3.2 Why the geographic criterion will not work

The geographic isolation criterion has a clear motivation as well. We have already considered a population that is composed of individuals on different continents; clearly organisms that are too

far away from each cannot be in the same population. Behind the geographic isolation criterion once again lies the requirement that members of populations be capable of some sort of relevant interaction: geographic barriers are important because they are barriers to some sort of relationship.

The geographic isolation criterion has a serious weakness: it is possible that populations should be treated as distinct for the purposes of selection theory, despite the fact that there is no geographic barrier between them. A failure of the right sort of interaction for delimiting populations may occur between two groups of entities for reasons other than the lack of time/space overlap between them. Indeed, in most places on Earth there are great many groups of organisms that must be treated as members of distinct populations for the purposes of deploying selection theory: African lions and African dung beetles do not form a single population for the purposes of selection theory. On its own, the notion of geographic isolation can do very little work in delimiting populations; it functions best as a complement to other ways of drawing boundaries around populations, such as the criterion that population members be conspecifics.

4.3.3 Why taxonomic criteria will not work

In the face of the difficulties with the geographic criterion just rehearsed, it is tempting to impose the requirement that members of a population be of the same species. I noted earlier that Lewontin has written that conspecificity is a requirement of selection (1978, 220). Though most selection theory models are deployed on entities that are members of the same species, it would be wise to avoid deploying the term "species" in selection theory if only because of the well-known indeterminacy that attaches to that notion. Besides, Darwin rightly considered the displacement of a local species by a better-adapted newcomer to the region an instance of

selection (1988[1859], 81); for contemporary examples of selection between species, see Human and Gordon (1996) and Rosenzweig (2003, ch. 6).

4.4 COMPETITIVE INTERACTIONS

John Damuth provides an interesting discussion of the notion of population in an article concerned to argue that clades do not function well as higher-level units of selection. As already noted, he conceives of the point of selection theory in the same way as I do (Damuth 1985, 1134), and argues in much the same way as I do, by reference to the point of the theory. Damuth is concerned to show that a clade is not a good population for the purposes of selection theory, and the reason he offers is particularly interesting:

> Consider a set of organisms, all of which are ultimately the descendants of a single arbitrarily chosen organism, including those that may have emigrated from the region occupied by the ancestor and are now dispersed throughout the species' range in various habitats. We do not consider this set of lineage members a biological "population," because at any given time its members are not necessarily functioning as economic interactants in a common milieu. (Damuth 1985, 1133)

The idea that members of a population interact in a fashion similar to that of economic competitors has other backers. Ghiselin requires that candidate populations for selection theory be circumscribed in terms of competitive relationships between them, relationships that parallel the sort of competition found in the sphere of economics (Ghiselin 1974, 51). Recall, too, that Darwin's notion of the struggle for existence is related to the economist Malthus' discussion of the struggle among people to earn a living in a competitive economy. Lennox and Wilson's insistence that the struggle for existence is necessary for selection invites a similar constraint on the determination of population size (1994). The analogy with economics is insightful; my

concern will be to state slightly more precisely what sort of competitive interaction is necessary among entities for them to be correctly grouped together into a single population in biology.

I turn now to my proposal for delimiting populations. My idea is to delineate populations on the basis of competitive relationships. I contend that the kind of interaction that is relevant for circumscribing populations is *competitive causal interaction*. An interaction is competitive if and only if it involves two entities each of which cause the production of their own descendants by means of some activity that also inhibits the descendant production of the other entity. Recall that I incorporate competition of this sort into my statement of the entrance rule for selection theory: selection theory applies to competitors. The systems over which we deploy the theory are systems of competitors that compete with each other. What the population criterion adds is a means for determining *how many* and *which* competitors form part of a single population.

There are a few reasons I talk about descendant production, rather than reproduction, in my definition of competitive causal relationships. Reproduction is a term usually reserved for organisms, and not all competitors are organisms. Anyway, it sounds a little odd to talk about alleles, not to mention chromatin marking schemes, as engaged in reproduction, as opposed to producing descendants. Another reason to talk about descendant production rather than reproduction is that I do not require any material overlap between parent and offspring (see Godfrey Smith 2000).

Most importantly, however, ancestor-descendant relationships are multi-generational; we are each descendants of our grandparents and, indeed, their grandparents too. Because two entities must inhibit each others' *descendant* production, rather than each others' *offspring* production, competitors may count as members of the same population because of what their *descendants* do. Competition between one competitor's grandparents and another competitor's

grandparents may continue in the form of competition between their descendants. This feature of the definition is what allows it to function as a criterion that picks out systems that endure through time.

This feature of the definition is actually critical for another reason. Population geneticists often consider populations with sub-groups in which cotemporaneous population members do not compete with one another. Equally, they consider discrete generation models in which none of the individuals from one generation exist in the next. That the descendants of cotemporaneous members of populations with both these may compete provides a mechanism by which individuals who are treated in population genetics models as part of the same population are counted as such according to my definition.

Looking back in time, the population will stretch back to the last two rivals that are the parents of every contemporary population member, for they are the ones who continue to compete by way of their descendants. This means that a population of alleles can stretch no further back than its coalescent, the allele (or more generally competitor) that is the ancestor of all contemporary alleles in the population. Indeed, the population will stretch back to the coalescent if the coalescent produces a mutant descendant whose descendants survive in the contemporary population; otherwise, it will stretch back to the first existence of multiple rivals that have descendants in the contemporary population.

4.4.1 Deploying the competition criterion

What's perhaps most interesting about the competition criterion is that it allows one to circumscribe a population of competitors by first picking out some initial member or members and then using the notion of competition to ascertain the scope of the broader population in which they are found. In some cases, one initial population member will be enough, though in

sexual populations, one member of each sex may need to be chosen as founding population members before the competition criterion can be deployed. Obviously, these individuals will have to be at least candidate reproductive partners for one another, though they may end up breeding with other population members that are added according to the criterion below.

Beginning with our founding population members, we require that each new candidate member of a population engage in mutual competition with at least one individual who is already a member of the population, though not necessarily an individual of a different type. This allows population members that never causally influence one another's reproduction to nevertheless be counted as members of the same population. For instance, in a population of twenty-six competitors, where each is named for each letter of the alphabet, A need not compete directly with Z to be grouped in the same population as Z. Rather, A need only compete with B who competes with C, who competes with D, …, who competes with Y, who competes with Z. Two entities belong in different populations if they do not compete with one another either directly or indirectly through intermediaries.

The competition that is essential to population circumscription is reflected in population genetics calculations insofar as the number of offspring that different types contribute to the next generation is proportional to the absolute number of offspring produced in the entire population. This fact in turn lies behind the standard use of *relative* frequency variables and *relative* fitness parameters in population genetics (Rice 2004, 10). Population genetics theory implies that the reproductive success of individuals of one type requires a corresponding failure on the part of other types, and competition is the causal mechanism that I suggest underwrites this relationship.

Lastly, the connection with economics should also be clear. Firms compete with other firms by doing what it takes to make more profit for themselves. In a free market, the excellence of one firm in profit-making imperils rival firms that accordingly lose market share.

4.4.2 Beneficial interactions

That members of populations may have non-competitive, even mutually beneficial interactions might seem to threaten the definition of population proposed here. Indeed, the reverse of inhibitory relationships of descendant production is actually put in play by competitive interactions: The inhibition of descendant production between A and B will benefit other competitors, such as C. But that A and B benefit C in this way need not imply that C should be placed in a different population from A and B. A or B may still have an inhibitory causal influence on the descendant production of C despite having promoting C's descendant production too. These inhibitory effects may be direct or indirect (perhaps mediated by a fourth population member, D), and they may exist in addition to the beneficial causal relations.

Similarly, "altruistic" individuals of the sort often modeled as members of populations that form sub-groups (e.g., Sober and Wilson 1998) can be admitted into populations even though they promote the reproduction of fellow population members because, in addition to their altruistic activities, they also do things that inhibit their fellows' reproduction. Any population that is not reproducing without bound is one in which some members are accumulating the resources necessary for reproduction at the expense of others who accordingly produce fewer descendants. Provided "altruists" do things like accumulate resources and thereby inhibit the reproduction of rival population members, they can be admitted into a population despite doing *other* things to promote their fellows' reproduction as well.

4.4.3 Recovering textbook characterizations of "population" as operationalizations

I noted earlier that the characterizations of population that I criticized above as inadequate conceptual definitions of population still work well as operationalizations for delineating populations. That's a good thing, since the criterion I offer is hardly deployable directly. Indeed, insofar as the criterion I offer makes critical use of causal vocabulary, and causal facts are not observable, the criterion must be deployed indirectly. Standard characterizations of "population" provide means to make such assessments.

Consider the traditional criteria in turn. Interbreeding is surely a sufficient condition for competition, for organisms that interbreed will typically *compete* for mates. They will also almost certainly share a niche. Accordingly, they can be expected to adversely affect each others' descendant production and benefit from doing so. Contiguity in space and time can also be used as a cue to delimiting populations, for such closeness is surely a necessary condition for competition. And, of course, members of the same species will usually share the same food supply, mates, predators, and other hazards when they occupy the same ecological range. Sameness of species membership will thus provide a reasonable cue for determining population membership too.

4.5 CRITICISMS OF THE COMPETITION CRITERION

In this section I consider some objections to the competition criterion I offer for forming populations out of competitors. Specifically, I consider both the possibility that it may be too strong and too weak.

4.5.1 The competition requirement is not too weak

The criterion I offer for picking out populations can be motivated by appeal to the failure of strategies that fail to abide by it. Suppose that one failed to follow the approach just rehearsed for delimiting populations by allowing there to be competition between members of a population and other individuals that are *not* considered members of the same population. Such competitive interactions between population members and non-population members could matter to the evolutionary trajectory of the population under consideration, and are almost certain to do so if the individuals with whom the excluded competitors compete are predominantly of one of several rival types. A failure to take these causal influences into account could lead to the generation of false predictions or unlicensed explanations of the population's dynamics, since causal influences that could bias relative descendant production rates within the population are being systematically ignored. Hence, the rivals who were left out of the population should be included in order to get the causal dynamics right.

That last argument merely shows that the influence of the rivals falling outside the population must be handled somehow, but it does not show that this influence must be handled by including them in the population of interest. Could we not handle their influence as akin to any other sort of environment influence? Such a treatment is perhaps possible, but several considerations speak against it as an option. The first is that competitors falling outside the population form a particularly impermanent sort of influence, since they will usually die off at roughly the same rate as the population members. Should one count the influence of their descendants upon the descendants of members of the population as a continuation of the same sort of external influence present in the initial generation? If one did so, then the influence of such external competitors would be contingent upon evolutionary change both among them and

within the target population. This would put the theorist in the business of monitoring evolutionary change in the competitors treated as external to the population anyway. And how would one tell the descendants of outsiders from the descendants of population members, especially if there were migration between these groups or worse interbreeding between population members and the individuals falling outside the population?

Failing to treat individuals that would be counted as population members according to my criterion as members of the population and instead treating them as a kind of environmental influence brings with it a host of headaches that can be avoided by expanding the population to include them. If we are to be in the business of explaining the spread of competitors and their adaptations, we might as well do so using a set of rules that makes such explanations tractable. Besides, we have variable selection models to handle environmental heterogeneity should the population picked out in the way I suggest span multiple niches. So, even if enlarging our population makes things more complicated because it means we deal with a population with multiple sub-environments, we know how to do that in at least some cases.

4.5.2 The competition criterion is not too strong

The criterion for circumscribing populations that I have offered might also be too strong and allow too few competitors to form part of the same population. To counter this criticism, I draw attention to the special nature of the causal relationship that must exist between a population member and at least one other population member. The relationship is no ordinary causal relationship. Surely, lots of biological entities have causal effects on one another. Many have adverse causal effects on one another's descendant production. But to be counted as members of a population, candidate individuals must each take actions that cause their own descendant production and also inhibit the descendant production of another population member. So, for instance, my criterion is much stronger than that used by Sober and Wilson to pick out "groups" on the basis of fitness-affecting interactions (1998, 92).

To see how the strength of the criterion eliminates potential counterexamples, note that should an allele be picked out as a founding population member, it will compete only with different alleles that contend for the same place on the genome. Even though alleles have fitness-affecting interactions with alleles elsewhere on the genome, they do not compete with them in sense I define. The detrimental effects that an allele has on other alleles are equally detrimental for itself.

The most worrying case for the criterion I offer is posed by groups of organisms that would seem to compete, but only weakly. For instance, warthogs and zebras both suffer from predation by lions, but they clearly should not be placed in the same population and treated as variants competing with each other. However, I have picked out predator avoidance as a mechanism that institutes competition: at least sometimes, the avoidance of a predator by one population member will lead the predator to pursue and even catch another population member.

Similarly, though I have picked out resource consumption as a mechanism for competition, it seems equally a mistake to place in the same population organisms whose resource consumption overlaps, but only a little.

In response to this problem, the zebras/warthogs problem, I claim that competition between types does not necessarily result from their sharing resources or from their being subject to predation by the same predators. My criterion requires that predation by lions on both warthogs and zebras would be grounds for considering them competitors in a single population only if, by escaping the lions, the zebras inhibit the reproduction of the warthogs, and vice-versa. However, the fact that lions prey upon both zebras and warthogs does not necessarily mean that this last condition is fulfilled. As I will argue, the condition will not be fulfilled unless the lions are serving to regulate the number of offspring that both the warthogs and the zebras have.

Warthogs and zebras both produce more offspring than can survive to maturity; their ecological environments contain many mortal factors that reduce a large number of juveniles to a smaller sample of reproducing adults. But the removal of some mortal influence that kills zebras and warthogs need not lead both the warthogs and zebras to have more offspring than they otherwise would. The slack created by the removal of one mortal cause might well be taken up by another one. Indeed, *diversity* in the causal influences on the size of a population does not imply a small overall population size due to stiffer population regulation.

Some natural populations will be such that their size is restricted by a single causal influence, such as the availability of nesting sites. Thick-billed parrots in Mexico are like this (Lanning and Shiflett 1983). Still, all sorts of mortal causal influences arising from the ecological environment may kill off the parrots. Raptors eat them, for instance, yet there is no evidence that the raptors serve to regulate the overall population size of the parrots. Generally, populations will be beset by a greater variety of mortal causal influences than just those that serve to regulate their sizes.

If, in our imaginary scenario, the lions are not serving to regulate the number of offspring that the zebras and warthogs have, just like the raptors are not serving to regulate the number of offspring the parrots have, then the avoidance of predation by warthogs does not lead to more zebras deaths than would have occurred anyway, and vice versa. Accordingly, the condition for treating the zebras and warthogs as members of the same population does not hold, for predator avoidance on the part of the zebras does not inhibit the reproduction of the warthogs, or vice versa. Much the same result will be obtained if we consider the case of partial overlap in resources. Members of two different populations can *consume* the same resources, without *depending* on the same resources.

Perhaps somewhat surprisingly, though predation by lions on the zebras and warthogs does not institute grounds for uniting them in a single population, if we have other grounds for treating the zebras, say, as a forming a single population, then predation by lions can lead to selection within the zebra population. A causal influence need not regulate the size of a population in order to institute selection within it. Returning to the thick-billed parrot case, it is plausible that genes that decrease parrots' probability of being eaten by a raptor will spread in the population by selection because they will increase the probability that their bearers get access to

one of the limited number of nesting sites. Predation by raptors can lead to selection among parrots. What it cannot do is justify grouping the parrots into a larger population with other birds that are also preyed upon by the same raptors.

Of course, I have assumed in my zebra/warthog thought experiment that lions are not serving to regulate both populations. I don't know enough about zebras and warthogs to affirm this as an empirical fact. Perhaps lions, or indeed some other causal influences, do serve to bind the zebras and warthogs together into a single population. But that fact would not on its own serve to undermine the definition of population I offer. Rather, in such a case, it would make sense to think of these two species as competing, as constituting rival types competitors, and hence as united into a single population.

I have already argued that members of different species can act as rival competitors, so I am committed already to viewing at least some instances of species-to-species conflict as ones governed by selection theory. Were it the case that selection was really going on in the zebra-and-warthog population between the zebras and warthogs, then we would expect to be able to draw inferences about the dynamics of that population using selection-theory models. Only a demonstration of such competition, coupled with a demonstration that the dynamics of the zebra-and-warthog population were not those that we infer using selection theory, would undermine the account of population offered here.

By deploying the above definition, we put ourselves in the position to group competitors into populations. It may sometimes be possible to spot competitors without it being possible to group them according to the above definition. In such a case, one can proceed no further with selection theory. But in cases in which one can group competitors into populations according to

the above definition, one has put oneself into a position to deploy the rest of the framework of selection theory as it is presented here. I turn to discuss that framework next.

5.0 CONCEPTUAL RESOURCES FOR THE CAUSE-TO-MODEL ALGORITHM

At this point, two steps toward the explication of selection theory sought here have been taken. I have put forward an entrance rule for selection theory, a rule that divides entities into two piles, those over which the theory may be deployed and those over which it should not be deployed. This initial step is crucial because it is not the case that all systems will behave as do ones undergoing selection; if the inferential rules of selection theory are going to permit the drawing of true conclusions, their applicability must be restricted to a subset of worldly systems. The right sorts of systems over which to deploy selection theory are competitors, as these are defined in chapter 3.

The second step already taken in the explication of selection theory is the specification of a rule that articulates how to determine the scope of the populations in which competitors are found. Populations are formed from competitors that compete with each other. That rule circumscribes populations without regard to the influences that the causal contexts of competitors have upon their descendant production. So on the view undertaken here there is nothing illegitimate about populations that span multiple ecological sub-environments, ones that are "heterogeneous" in Brandon's sense (1990). Because populations may be causally different from one another, competitors in different populations can be expected to differ in the sorts of dynamics they exhibit because the causal influences operating on them vary. Classical population genetics exhibits a corresponding variety of equations that apply to this diversity of populations.

The next three chapters are dedicated to exhibiting the connection between the variations in causal influences operative in populations and the various equations of classical population genetics that capture the dynamics of those populations. The next chapter consists in an algorithm for generating deterministic, "Wright-Fisher," classical population genetics equations from causal information about populations. The chapter that follows that one is dedicated toward showing how to use deterministic models as contributors to more general stochastic models of population dynamics. Making the cause-to-model connection takes considerable preparation, however, and the current chapter does the definitional and explicative work necessary to make the algorithm comprehensible.

That preparatory work, in outline, consists in the following tasks. The causal influences operative in a population must be categorized along one set of dimensions and then categorized along another set of dimensions before they can be fed into the algorithm. The first categorization scheme is counterintuitive, too, and must be motivated. Furthermore, the various sorts of entities to which population genetics equations make reference must be defined. While we have a definition already of "competitor," most population genetics models feature relative frequency terms that refer to other sorts of entities, including gametes, zygotes, mating pairs, sub-groups, and substructures. So this chapter prepares the way for the algorithm offered in the next chapter by undertaking to define and motivate some causal notions, as well as to offer a principled account of the entities and groupings of competitors that appear the in population genetics formalism.

However, before I get to those matters, I want to give the reader a general sense of what the algorithm we will be working towards looks like. I also want to give the reader a sense of why an algorithmic approach to generating classical population genetics equations from causal

information is sensible. So the next section offers a quick preview of how the algorithm works. Knowing where we are headed will perhaps help make sense of why we make the distinctions and undertake the definitions that we do in the body of this chapter.

5.1 A PREVIEW OF THE ALGORITHM

The algorithm I offer follows a *constructive* approach to modeling population dynamics. The driving idea is to present a function for determining, for any given population, an appropriate classical population genetics equation that is no more mathematically complex than it needs to be to make it possible to accurately model the dynamics of the target system. This means that the equations yielded by the algorithm never feature variables that get set to null values such that the variables could be dropped from the equations without this impacting what sorts of dynamics are inferred from the equations. That is, the equations I generate are never too complex for the population at hand, such that they include variables that could represent causal influences were they to operate, but which must be set to values that void their inferential import.

Because I seek this sort of minimal complexity in the models generated by the algorithm, I do not restrict myself to the consideration of equations that are all versions of a single equation. I do not begin with a maximally general equation and then show how to set its variables at null values to generate simpler equations with narrower scope for populations beset by fewer and less various causal influences. Instead, I begin with a decision-tree that determines which of a variety of types of basic equation is appropriate for some system. So, for instance, whether the rival alleles that have been picked out as the target competitors in some population interact with genetic variations elsewhere on the genome is a criterion used to determine whether or not a multi-locus model, rather than a single-locus model, is appropriate for that population of

competitors. Similarly, querying whether alleles undergo gene-by-sex interactions is used to determine whether equations characteristic of sex-dependent selection are appropriate, or whether a simpler model in which sex differences are ignored can be used instead. Ultimately, six different types of fundamental model are distinguished by means of a decision tree, ones that lead to six different sorts of recursive equations that are used to capture the dynamics of the systems. I give a set of rules for choosing whether to pursue a model that treats a system using recursions on gamete frequencies, zygote frequencies, or sex-dependent versions of these. I also say under what circumstances haploid and haplo-diploid models are appropriate.

While the decision tree just discussed yields an understanding of *what sort* of model to use for a population, it does not yet yield an understanding of *how many variant types* we must consider in that sort of recursive equation. So the second step of the algorithm involves composing a directed acyclic causal graph showing transitions between different stages in the lifecycle of the entities in the target system, one that yields simple equations that feature appropriate relative frequency terms. Classical population genetics equations feature four sorts of relative frequency terms, terms for haploid frequencies, gamete frequencies, zygote frequencies, and mating pair frequencies. For ease of discussion, I will call haploids, gametes, zygotes, and mating pairs, *individuals* in what follows.[16] Equations that are recursions on gamete frequencies vary; they may include relative frequency terms for two, three, or more variant gametes, and these may turn up in multiple substructures. The whole lifecycle of the population will be sensitive to how many variant gametes there are in the population: more variant gametes implies

[16] I define "individual" later on; indeed, it is one of the main endeavors of this chapter to do so. Note that allele frequencies are not featured in classical population genetics models used to capture population dynamics. Rather, in single-locus models, gametes are distinguished by a single genetic variation, leading some writers to refer casually to gamete frequency variables as allele frequency variables. But such models are special cases; multi-locus selection models, for instance, do not feature any variables that could possibly refer to allele frequencies. To achieve the sort of generality pursued here, we must maintain talk of gamete, zygote, and mating pair frequencies when discussing diploid selection.

more variant zygotes, more variant zygotes implies more variant mating pairs. By graphing each step of the lifecycle according to a fixed set of instructions, we generate equations featuring the right number of relative frequency terms referring to the right number of individuals, where these differentiated in terms of the alleles that they bear. The graphs have other uses too; edges are weighted by coefficients in the graph to represent biases in how gametes, zygotes, and mating pairs form and produce each other as the lifecycle progresses. Assortative mating parameters are a paradigm instance of parameters that measure such biases.

Below is an example of a lifecycle graph, one appropriate for competitors that are rival alleles at a single-locus in diploid organisms undergoing sex-dependent selection. The diamonds represent gametes and the rectangles represent zygotes. The one-half coefficients in the second graph represent meiosis. The gametes are distinguished using a three-part indexing system, in which each index is separated by a comma. I will fully explain it later on. In this graph, the third index can be ignored, the second index refers to sex, and the first index refers to the alleles born at a single genetic locus:

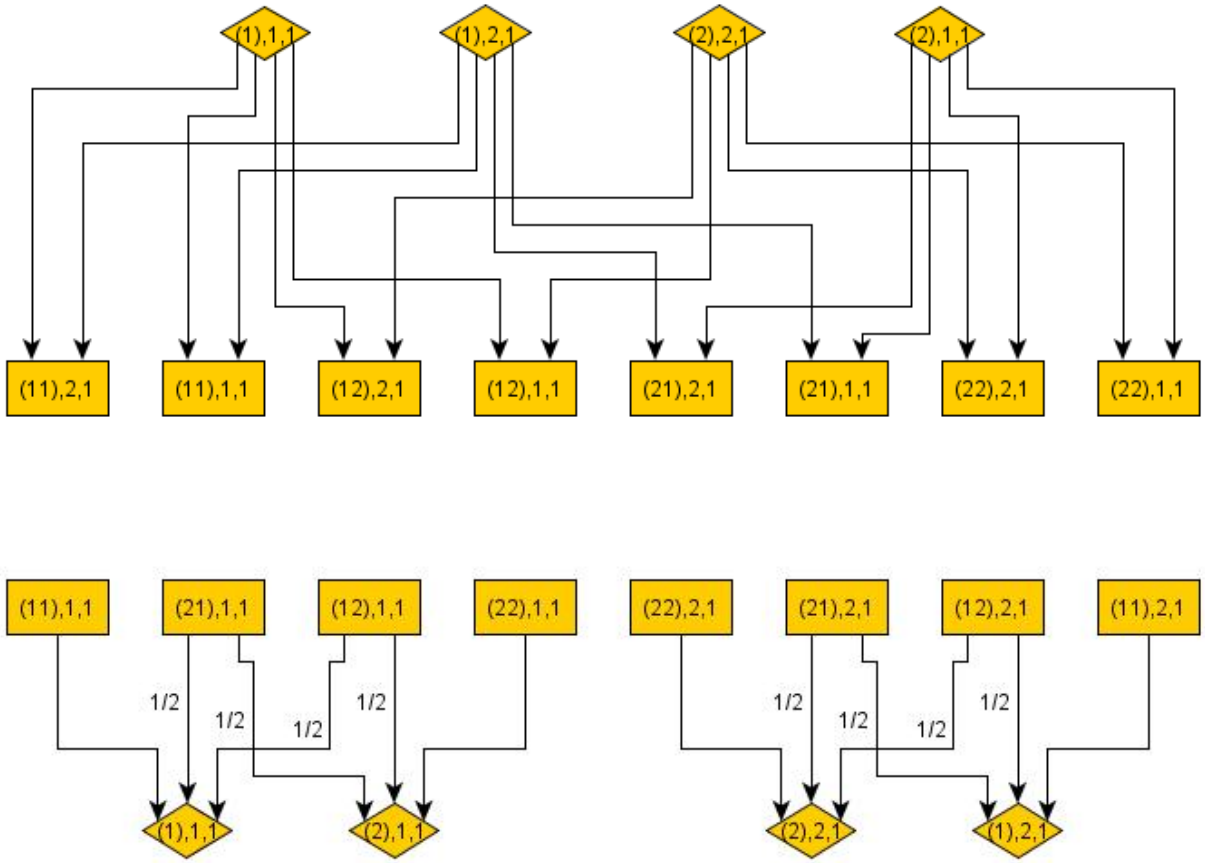


Figure 5.1: A lifecycle graph for a population undergoing sex-dependent selection

The graphs for the lifecycle yield either one or two systems of fundamental equations which represent the lifecycle of our population; the above lifecycle graph yields two systems of equations. The values of variables associated with descendant nodes are fixed mathematical functions of the values of variables associated with ancestor nodes. All populations, with the exception of haploids, will go through two lifecycle stages and these must be graphed separately, as above, such that two systems of equations are generated.

The equations yielded by the graphs are not sufficiently complex, however, to capture the dynamics of the target populations, nor are they intended to be. Indeed, the equations yielded by the graphs include no viability selection coefficients at all, the coefficients that represent "selection" in its most basic form. However, the equations are all versions of more complex equations that are generalizations of them, ones that *do* include such things as relative fitness coefficients. So, in the third step of the algorithm, I take advantage of the possibility of generalizing equations by substitution to yield more complex models. Indeed, the guiding idea behind the algorithm has been to rely upon this third substitution stage as much as possible.

To illustrate the relationship that models have when one is a version of another, such that the more general one can be generated by following a substitution rule for transforming the more specific one, consider two equations that are components of systems of equations used in two-locus Wright-Fisher models that are recursions on gamete frequencies. The first equation is drawn from a system of equations that governs the evolution of a population in gametic disequilibrium; it captures the next generation gamete frequency of A_1B_1 gametes this way (Rice 2004, 41-45):

$$\frac{x_{11}^{'} = x_{11}^2 + x_{11}x_{12} + x_{11}x_{21} + (1-r)x_{11}x_{22} + rx_{12}x_{21}}{\overline{w}}$$

Equation 5.1

The gamete frequencies are all picked out by *x* variables, with the first subscript picking out which of two alleles is borne at the first locus, while the second subscript picks out which of two alleles is borne at a second locus. If we add selection into the picture[17], we get the following set of equations:

$$x_{11}^{'} = \frac{w_{1111}x_{11}x_{11} + w_{1112}x_{11}x_{12} + w_{1121}x_{11}x_{21} + w_{1122}(1-r)x_{11}x_{22} + w_{1221}rx_{12}x_{21}}{\overline{w}}$$

Equation 5.2

The first model is a version of the second insofar as setting the values for each of the w_i variables in the second set of equations to 1 yields the first set of equations.

Whenever values can be deployed in the place of variables, such that a determinate version of a more determinable equation is reached, the opposite process can occur too, such that a determinable generalization of a determinate equation is generated. Any term in an equation can be regarded as having an implicit coefficient of "1" associated with it, and that implicit coefficient can be replaced by a variable, or even more generally a function, so as to generalize the equation in which the term with the implicit coefficient of "1" is found. This sort of generalizing substitution, in which a more general term replaces a more specific one, is a legitimate sort of substitution for predicates in our language quite generally (Brandom 1998, ch. 6).

For example, if for every x_yx_z on the right-hand side of (1) above, we make the following substitutions:

$$x_yx_z \rightarrow w_{yz}x_yx_z$$

Equation 5.3

[17] In the technical language that is to come, "adding selection" amounts to adding variables such that the impact of a pervasive interactive ecological causal influences can be represented in the model.

in which the replaced terms is on the left of the arrow and the replacing term on the right, we can generate (2). This process of substituting variables into basic equations to replace implicit unitary coefficients is the sort of process used in the third step of the algorithm, and it is used to accommodate all the causal influences that were not accommodated in the previous stages.

The final stage of the algorithm involves the collapse of the two systems of equations that are generated at the graphing stage and manipulated at the substitution stage. For instance, when considering a simple population, such as one whose individuals are differentiated by their bearing one of two alleles at a single locus of interest, we would find ourselves with two systems of equations at the end of the substitution stage. This first system of equations is inferred from the lifecycle graph for the gamete-to-zygote portion of the lifecycle, and here we are imagining that the individuals were not subject to any causal influences that would have had to be represented through generalizing substitutions:

$$p_{11} = \frac{p_1 \times p_1}{p_1 + p_2}$$
$$p_{12} = \frac{p_1 \times p_2}{p_1 + p_2}$$
$$p_{21} = \frac{p_1 \times p_2}{p_1 + p_2}$$
$$p_{22} = \frac{p_2 \times p_2}{p_1 + p_2}$$

Equations 5.4

where p_1 refers to a gamete bearing the A_1 allele, p_2 refers to a gamete bearing the A_2 allele, p_{11} refers to a homozygote for the A_1 allele, p_{12} refers to a heterozygote, p_{21} refers to a heterozygote, p_{22} refers to a homozygote for the A_2 allele.[18] The second system of equations is inferred from the lifecycle graph representing the production of next generation gametes from this generation

[18] The reason why we have two terms for heterozygotes is discussed in the algorithm proper, where a more complicated indexing system is used.

zygotes, and these have transformed by generalizing substitutions. The zygote frequencies have been weighted by relative fitness coefficients and average fitness parameters:

$$p_1' = \frac{w_{11}p_{11}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}} + \frac{1}{2}\frac{w_{12}p_{12}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}} + \frac{1}{2}\frac{w_{21}p_{21}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}}$$

$$p_2' = \frac{w_{22}p_{22}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}} + \frac{1}{2}\frac{w_{12}p_{12}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}} + \frac{1}{2}\frac{w_{21}p_{21}}{w_{11}p_{11}+w_{12}p_{12}+w_{21}p_{21}+w_{22}p_{22}}$$

Equations 5.5

When these two systems of equations are collapsed, we eliminate the zygote frequency variables on the right-hand side of the second set of equations by replacing them with the right-hand sides of the first system of equations that are set equal to them. This yields a single system of equations that function as recursions on gamete frequencies.[19] Collapsing the system of equations by replacing each zygote frequency in the second system with the function that specifies it from the first system of equations, and doing a little mathematical simplification, we arrive at the following set of equations:

$$p_1' = \frac{w_{11}p_1^2 + w_{12}p_1p_2}{w_{11}p_1^2 + 2w_{12}p_1p_2 + w_{22}p_2^2}$$

$$p_2' = \frac{w_{22}p_{22} + w_{12}p_1p_2}{w_{11}p_1^2 + 2w_{12}p_1p_2 + w_{22}p_2^2}$$

Equations 5.6

These may be recognizable as the standard single-locus diploid selection model.

Note how in the last system of equations above, relative frequency variables for zygotes do not appear, having been replaced by the right-hand sides of the first system of equations, generating a recursive system of equations. The reason why we must keep the systems of equations apart until after the substitution phase of the algorithm is so that we can deploy

[19] All Wright-Fisher systems of equations are recursions on one type of individual or another; gamete recursions are the most analytically tractable.

substitution rules on zygote frequencies, something that requires having equations that bear zygote frequency terms. We cannot weight the relative frequencies for zygotes in a system of equations that does not feature zygote frequency variables; at least, we cannot do so by following general rules for making substitutions that reflect causal influences over individuals, no matter what sort of individual we consider. Hence, in the algorithm, we put off the sort of collapse of equations just illustrated until everything else is done.

To close this preview section, I trumpet the constructive character of the algorithm. Since at each stage the algorithm guides decision-making about how to generate or modify equations to capture the dynamics of the system using information about the action of causal influences upon it, the result is that the equation includes no more complexity than is required to capture the dynamics of the system. We do not make a generalizing substitution, for instance, unless the population is beset by causal influences whose impact must be captured by the variables that we would introduce by making the substitution.

The use of a constructive approach is no guarantee, of course, that an analytically tractable classical population genetics equation will be yielded by the algorithm, but when such an equation *can* be deployed to capture the dynamics of the system, the algorithm will yield that equation rather than a too cumbersome, too general alternative, something that might not be analytically tractable. Recall that the point of population genetics is to demystify adaptation, polymorphism, and altruism. Analytically tractable equations can be used to *prove* that a system will become better adapted, or remain polymorphic, or consist in part of altruists. Such proofs go a long way to showing how adaptation, polymorphism, and altruism could be explained using selection theory.

5.1.1 Minimally complex population genetics equations are inferred

While an algorithm for producing classical population genetics models sounds daunting, and what I have to offer is not especially simple, it is not so strange that it should be possible to *say* under what conditions minimally complex models are applicable. Population geneticists have long been able to understand, at least implicitly, under what conditions analytically tractable models are appropriate for populations of interest. We see even the simplest basic types of models, single-locus models, being deployed for populations in nature (e.g., Hori 1993; Schemske and Bierzychudek 2001; Smith 1993; Stahl et al. 1999). Given that it is possible to recognize that a single-locus model rather than a more complex multi-locus model may be deployed over some system, it may be possible to make that sort of understanding *explicit* through a rule. The point holds generally: if it possible to recognize that it is appropriate to deploy a frequency-dependent selection model for some system (e.g., Hori 1993; Gigord, Macnair, and Smithson 2001), then presumably it is possible to say under what conditions it is appropriate to do so; if it is possible to recognize that a temporally variable selection model is appropriate for some system (e.g., Turelli, Schemske, and Bierzychudek 2001), then presumably it possible to say when one should deploy that sort of model too, and so on.

One might wonder whether that last argument is too strong. There are surely *some* fragments of language, paradigmatically observation concepts, whose applicability we can recognize without doing so on the basis of explicit rules. We can recognize when it is appropriate to call something red, but few of us can say under what circumstances it is appropriate to call something red. Generally, there must be language fragments for which it is not possible to say under what conditions it is appropriate to deploy them, at least in a non-circular fashion. However, classical population genetics models are probably not among these.

The applications of population genetics to natural populations cited above show that researchers affirm the (typically approximate) appropriateness of some formal mathematical model for their target populations only after extensive research designed to establish what are the causal influences of the genetic variations that they have found in their populations. The applicability of a classical population genetics model for some system is *inferred* on the basis of an understanding of the causes to which the population members are subject, something that itself is affirmed on the basis of empirical evidence. Endler has dubbed this the "functional approach" to selection:

> It is an excellent trend that more and more people are working on the function and ecology of adaptive traits because the functional approach allows one to predict variation in fitness and evolutionary change rather than simply proving its existence. A particular advantage of knowing the function of a trait in detail is that is allows specific predictions about fitness and the direction of evolution. (2000, 253-54)

Since it possible to deploy models of population dynamics on the basis of the function and ecology of adaptive traits, it might well be possible to understand explicitly when it is appropriate to deploy which models on the basis of which features. Making explicit the inferential relationships between salient features of populations, here *causal* features, and dynamical models is what the algorithm is designed to do in the general case.

5.1.2 The scope of the algorithm

The algorithm I offer has a less general scope than does selection theory as a whole. Specifically, the algorithm I offer shows how to generate causally interpretable Wright-Fisher equations, which are suitable for populations with discrete generations. Beyond their mathematical simplicity, the reason I choose this subset of classical models is that they involve a particularly strict basic assumption about non-overlapping generations, one that gives them an especially

flexible recursive structure to model a variety of causally distinct populations, including ones that have causally complex relationships to their contexts and to each other. As the algorithm makes clear, Wright-Fisher models can be used to address the impact of causal influences arising from alleles elsewhere on the genome and across chromosomal pairings, interactions between genotypes and sex differences, interactions between individuals and ecological causal influences, as well as interactions among the individuals within a population. I am pushing a causal interpretation of selection theory, and the crucible for such an interpretation is its adequacy as an interpretation of Wright-Fisher models because these models can represent a greater variety of causally complex situations than can be represented on alternative approaches, including models with overlapping generations, Cannings models, Moran models, and others. Despite its limitations, the algorithm I offer for connecting causal information about a population to a definite equation governing its dynamics covers a wide enough breadth of population genetics models that it should provide substantial evidence for the claim that selection theory, or at the very least classical population genetics, is causally interpretable.

Indeed, the claim that selection theory is causally interpretable is probably more interesting than the algorithm deployed to justify the assertion. The algorithm is a awkward, multi-step, cumbersome affair, of little use for the generation of interesting population genetics models, many of which have been developed already anyhow, and all of which could be generated intuitively by an able-minded population geneticist without the use of any sort of algorithm, much less an awkward one such as that which I have developed. Indeed, the algorithm makes explicit the norms operating behind the intuitive generation of population genetics models by population geneticists, norms I simply could not have learned without relying on textbook examples of the various sorts of models the algorithm generates. The algorithm does not provide

a way of generating novel models so much as it presents explicitly what has already been intuitively understood for quite some time by population geneticists who have been generating such models without any explicit tools.

What the algorithm does serve to do is provide a definite sense in which population genetics models can be causally interpreted. Whether or not population genetics is a causal theory has become a hotly contested issue in the philosophy of biology. The converse view that population genetics provides nothing more than a statistical summary of population dynamics has come to be known as the statisticalist view (Matthen and Ariew 2002, 2005; Walsh, Lewens, and Ariew 2002; Ariew and Lewontin 2004; Walsh 2004, 2007). An algorithm for generating population genetics models on the basis of causal information is a powerful counter to the statisticalist view; it is a causal interpretation of selection theory in a strong sense. The algorithm provides a means of using causal information as an *inferential basis* for population genetics models. Since population genetics models equations yield facts about the dynamics of the population to which they are applied, the models can be understood as providing a link between causes and dynamical effects, a link that can be understood as *explanatory* because of its reliance on causal information as a critical input. In short, I use a causal interpretation as an argument for causal interpretability.

5.2 CAUSAL CLASSIFICATION IN CLASSICAL POPULATION GENETICS

Officially, at this point, we have know how to pick out systems over which it is appropriate to deploy population genetics models, and we know how to group them into populations, but an account of which populations have their dynamics explained by which models has not yet been offered. In order to say how to model a given population, we have to know how to carve up the

causal influences acting on its members. The next section will be concerned with a two-step categorization scheme that yields a categorization of contextual causal influences suitable for use in the cause-to-model algorithm whose presentation follows. However, that categorization scheme functions in a peculiar way: I pick out contextual features of competitors' circumstances as causally responsible for the dynamics of competitors. This requires an understanding of what I dub the principle of causal inversion. I discuss this principle before I discuss how causes of population dynamics must be broken down for use in the cause-to-model algorithm.

5.2.1 The Inversion Principle

I simultaneously embrace two commitments that would seem to be in tension with one another: 1) a commitment to causally interpreting classical population genetics equations such that the key causal variables in those equations are relative frequency terms, paradigmatically picking out haploids, gametes, zygotes, and mating pairs; and 2) a commitment to generating many of these equations from an algorithm that takes as inputs causal facts about a population, but that, curiously, does not feature among those inputs facts about the causal influences of haploids, gametes, zygotes, and mating pairs.[20] The compatibility of these two commitments rests upon what I call the *inversion principle*.

The inversion principle is a corollary of Woodward's account of causality, though any account of causality that violates it is likely to be criticized for just that reason. The inversion principle states that whenever C is a condition for X to cause Y, C, too, is a cause of Y, provided X obtains. Essentially, the inversion principle allows us to invert the roles of condition and cause, hence the name I give it. The inversion principle is what allows one to generate classical population genetics equations using an algorithm that does pick out haploids, gametes, zygotes,

[20] The exceptional cases are models featuring frequency-dependent selection functions.

and mating pairs as causes, and yet to nevertheless causally interpret the equations that it yields such that these same individuals are regarded as causes of population dynamics.

The idea that we can exchange causes and conditions in accordance with the inversion principle has considerable intuitive appeal. In many cases, causes and contexts can be exchanged for one another without jarring our intuitions too much. Consider these examples: thirst, when one is offered potable water, causes drinking, so offering potable water to someone who is thirsty causes them to drink it; aviophobia, when on an airplane, causes fear, so riding on an airplane, when one has aviophobia, causes fear; consuming phenylalanine, when one has the gene for PKU, causes retardation, so the gene for PKU, among those who consume phenylalanine, causes retardation.

Population geneticists avail themselves of the inversion principle implicitly when modeling populations. Instead of introducing variables into population genetics equations to handle the influence of contextual causes, say an ecological causal influence that besets only some population members, population geneticists contextualize the types in their populations by placing them in distinct sub-environments, ones in which the causal influence is operative and ones where it is not. In doing this, they ascribe distinct relative fitness coefficients to the different types in the different sub-environments. The relative fitness values assigned to the different types in the different sub-environments then reflect the impact of the sub-environmental context because that context is a distinct condition in which the different types compete. By handling ecological influences in this way, population geneticists maintain the use of fitness values that quantify the causal influence of the individuals whose relative frequencies they weight. Yet, by placing these individuals in distinct contexts, where the contexts are carved out

in terms of the contextual causal influences operative within them, population genetics models deploy fitness coefficients that quantify the causal influence of sub-environments too.

Indeed, we can ask two sorts of "what would have been different questions" of the sort that Woodward argues bring out causal relationships (2003, 187-94). We can ask: What would have been different if this type A individual had developed in sub-environment 1 rather than sub-environment 2? By manipulating (if only conjecturally) the sub-environment in which the organism develops, we can bring out how "wiggling" the sub-environment of an individual produces causal consequences for the individual's development, and ultimately its fitness. But we can also ask what would have been different if an individual developing in sub-environment 1 had had the genotype characteristic of a type B zygote, rather than that of type A zygote. Again, by conjecturally manipulating the genotype of the organism, we can bring out how the genotype causally impacts to development and fitness.

The appropriate response to this situation, in which we have available two contrast classes, a contrast in genotype and a contrast in sub-environment location, is to take it that *both* the genotype and the sub-environment membership of an individual have a causal impact on its development and fitness. So, population geneticists introduce ecological causal influences into their models by placing individuals in distinct causal contexts, and assigning them distinct relative fitnesses in each. When they do this, however, they preserve the potential for causally interpreting the resultant equations: the sub-environmentally contextualized individuals remain causes of population dynamics in the equations, ones whose frequencies can be (at least conjecturally) manipulated to alter how the population evolves. What makes this dual interpretation possible is the inversion principle. When one is considering whether the genotypes of the individuals are a causal influence over population dynamics, one treats the sub-

environments as parts of the causal context. Similarly, when one is considering whether sub-environment is a causal influence, one treats the genotypes of the organisms as parts of the causal context in which different environmental factors have their influence.

The inversion principle does lead to some counter-intuitive consequences, especially with respect to preventative causation and causation by omission. An absence of antidote technically causes death by snakebite, at very least when one has just been bitten by a venomous snake, but even perhaps also when such snakes are simply nearby. The difficulties posed by causation by absence and preventative causation are discussed by Woodward in chapter 2 of *Making Things Happen* and there's little point rehearsing them here. It is worth noting, however, that one consequence of the fact that the inversion principle makes us treat absences as causes surfaces in population genetics models.

Contextual causal influences that beset only some population members inevitably produce *two* causal contexts, one formed of the individuals struck by the causal influence and one formed by the individuals who are not struck by the causal influence. Some varying contextual causal influences have natural correlatives that are themselves easily understood as causes; for instance moist soil conditions contrast with dry ones. Other contextual causal influences lack correlatives that are easy to think of as causes: toxins in one region of the ecosystem contrast simply with a lack of toxins in other regions. But no matter how easy to think of the correlative of a cause as itself a cause, two sub-environments are warranted whenever one considers a causal influence that affects only some population members, one sub-environment formed by the presence of the causal influence and another formed by its correlative, even if that correlative is just the absence of the cause that strikes elsewhere. An absence of toxins is just as much a contextual cause in population genetics modeling as is the presence of toxins for

populations that span regions of the ecosystem in which toxins are at work and regions in which they are absent.

5.2.2 Key Concepts of the First Categorization Scheme

Having established that there is nothing incompatible between, on the one hand, the use of an algorithm for the generation of population genetics models that takes as inputs only facts about contextual causes, and, on the other hand, causally interpreting the resulting equations such that variables not treated as causes in the first stage are interpret as causes in the questions, I now turn to discuss the breakdown of contextual causes that will be used in the algorithm. An initial dissection of causal influences along three axes is necessary. Causal influences may be either interactive or non-interactive, discriminate or indiscriminate, and pervasive or non-pervasive. I tackle each of these categories in turn.

5.2.2.1 Causal interaction. Cartwright (1979) offers a gripping example of an interactive causal influence. Imagine I drink an acid poison. Normally, this will kill me since drinking an acid poison causes death. However, if I have just drunk an alkali poison, drinking an acid poison will save my life. This means that the recent consumption of an alkali poison is an *interactive* causal influence upon acid poison intake. Drinking an acid poison has one sort of effect in one sort of context and another sort of effect in another context: in the context of no recent consumption of an alkali poison, drinking acid kills, while in the context of recent alkali poison consumption, drinking acid saves lives. The influence of any sort of cause can exhibit this sort of contextual dependence when combined with another causal factor.

Cartwright's acid-drinking scenario contains a nice symmetry. On the one hand, we can consider the acid poison to be the causal influence of interest and then consider its influence on survival in two distinct contexts. On the other hand, we can consider the consumption of the

alkali poison the prime causal variable of interest, one whose influence in contingent upon whether an acid poison will shortly be consumed or not. Both these approaches come to the same thing, a fact that is most easily expressed formally. Letting *s* = survival, *al* = the consumption of an alkali poison, *ac* = the consumption of an acid poison, and allowing all variables to take on values of one or negative one to respectively represent occurrence or non-occurrence, we can write

s = *ac* x *al*.

Equation 5.7

This expression says that drinking an acid poison will kill you unless you drink an alkali poison, too; we can also interpret it as saying that drinking an alkali poison will kill you unless you drink an acid poison. Indeed, it says both these things at once.

The notion of causal interaction at play in Cartwright's acid example is the one in use in what follows. Picking out causal influences as either interactive or not is critical to botanizing them such that we can make sense of which models, featuring what sorts of parameters, get deployed over which sorts of populations. When I pick out causes as interactive causal influences, I mean to pick out ones that affect different types in the population differently, depending on what type of individual the individuals are. Just as acid poisons have different influences on different types of people depending on whether they have recently drink alkali poisons, various contextual causal influences on individuals may affect them differently depending on their genotypes, or more generally what competitors they bear.

5.2.2.2 Pervasiveness. Another key concept in the breakdown of causal influences that forms the input for the cause-to-model function is *pervasiveness*. A causal influence is pervasive if it affects all members within some grouping; it is non-pervasive otherwise. We must model

contextual causal influences differently depending on whether they are pervasive or non-pervasive within such things as whole populations, distinct substructures, and distinct subgroups.

Surprisingly, understanding *non-pervasive* causal influences over individuals as real and legitimate causal influences over population dynamics comes more naturally than does understanding *pervasive* contextual causal influences in this way. If a population spans two sub-environments, say toxic and non-toxic soil, then it is relatively easy to understand how the relative fitness values of the different types in the different sub-environments reflect the causal impact of the soil. What's more, sub-environment membership has the usual implications of a causal relationship: *were* we to move plants from the toxic region to the non-toxic regions, then they would have more offspring on average. Understanding toxic soil as an inhibitor of plant growth is straightforward, especially when some members of the population grow in non-toxic soil and we can imagine moving plants from one region to another.

In mathematically more basic cases in which we are dealing with causal influences that are pervasive, it is more cognitively challenging to think of these causal influences responsible for impacting population dynamics. When soil conditions do not vary, when all the soil encountered by all the plants is non-toxic, for instance, it would seem odd to blame the dynamics of the population on a lack of soil toxicity. It seems odd to do this even if the genotypes of the plants would have different causal influences on their relative reproduction rates were they growing in toxic soil. Still, causal scenarios featuring organisms whose genotypes have different causal influences, causal influences that are contingent upon some invariant environmental parameter, are causal scenarios in which that invariant environmental parameter is a causal influence upon population dynamics.

That last fact is just a consequence of the inversion principle: If the plant genotypes have differing causal influences over reproduction only in the context of non-toxic soil, the non-toxicity of the soil is a cause of the differential reproduction rates of the plants. One of the linchpins of the causal analysis conducted here is a stubborn determination to think of population dynamics as resulting from contextual causal influences, rather than thinking of those dynamics as the result of the causal influences of differing types of individual featured in the model. The inversion principle makes either perspective legitimate, but the algorithm requires focusing on causal impact of context.

To see that it possible to trade in the type-centered perspective, according to it is the different causal influences of the different types that drive population dynamics, for the context perspective, which picks out the causal context as the driving force behind population dynamics, consider this imaginary scenario. Hawks with sharper vision react more often, and from greater distances, to the movement of prey and accordingly catch more prey, spend less time and energy hunting, are less likely to starve, and ultimately have more offspring than their less well-endowed counterparts. Indeed, it is plausible (though by no means certain!) that there was a time in the history of hawk lineages in which alleles spread throughout hawk populations because they had just this sort of beneficial causal influence on their bearers' long-range eyesight.

Deploying the inversion principle, we can characterize this same episode of evolution as resulting from ecological causal influences. Prey movement had a different influence on different types of ancestors of our contemporary hawks. At least some instances of motion on the part of prey animals caused hawks with some alleles to attack but would have failed to trigger the same reaction on the part of their counterparts bearing different genetic variations. The hawks more easily provoked to attack by movement on the part of more distant prey, prey so distant that their

counterparts would have overlooked its activity, were caused to reproduce by that prey movement. It is plausible that prey motion literally caused the differential reproduction of hawk alleles.

Either way of presenting the scenario just considered is legitimate, but the cause-to-model algorithm of the next chapter requires that we use the second one. Indeed, the hawk dynamics just imagined could be blamed on *any* feature of the causal context. While we could just as well blame the spread of hawk alleles for improved eyesight on prey motion as we could blame it on the alleles for better and worse eyesight borne by the hawks themselves, we could just as easily blame the evolutionary episode on the pervasive influence of sunlight, which is requisite for hawks to see at great distances. In the context of moving prey and hawk alleles for differing visual sensitivities, sunlight causes the spread of the alleles for sharper vision. Equally, in the context of sunlight and moving prey, the hawk alleles cause their differential reproduction, and in the context of sunlight and genetic variations, the prey movement causes the differential reproduction of hawk alleles. We can alternate between these three ways of presenting the hawk scenario because we can at least imagine how things would go differently in the hawk population were there no genetic variations, no motile prey, or no sunlight.

Indeed, we could even blame the evolution of hawk visual acuity on factors that are not ecological factors. Genes at other loci that are fixed throughout the population, for instance ones that are essential to the hawks' development of retinas, could be held responsible for the spread of other genetic variations in the hawks that matter to eyesight. In the context of differing alleles for better or worse vision, as well as sunlight and prey movement, genes that contribute to eye development cause the spread of the hawk alleles. Faced with this sort of embarrassment of riches we must simply pick out any pervasive feature of the causal context whose presence is

requisite for the rival types to have different causal influences and blame that feature of the causal context for the differential performance of the rival types. If we do so, we have what we need to use as an input for the cause-to-model algorithm. For convenience, I suggest we blame ecological factors, since that is typically what population geneticists have in mind when talking about the influence of "selection."

When we take the perspective I am encouraging, we can understand fitness coefficients, the *w* parameters, in equations such as the familiar diploid selection model:

$$p' = \frac{p^2 w_{11} + pqw_{12}}{p^2 w_{11} + 2pqw_{12} + q^2 w_{22}}$$

Equation 5.8

as quantifying the causal influence of the ecological environment on the different genotypes. They do this just as much as they quantify the causal influence of the distinct genotypes with which those coefficients are paired. The ecological environment simply does not vary in the above model (or better, the model is suitable for a population in which the ecological environment does not vary); only the alleles and the genotypes of the zygotes vary. But the ecological environment is no less a cause of population dynamics just because it affects all the zygotes.

Generally, the difference between population genetics models that feature distinct sub-environments, and ones that lack distinct sub-environments, is *not* that contextual causes are at work in the former but not the latter. Rather, in the models featuring individuals that are not placed in distinct sub-environments, the contextual causes that matter to the dynamics of the population are *pervasive*; the contextual causes of the sub-environments affect *all* the individuals in some grouping in the population.

5.2.2.3 Discrimination. So far, we have considered two key concepts that we will use to carve up causal influences for use in the cause-to-model algorithm, interactivity and pervasiveness. The last key concept we need to complete the first categorization scheme is that of discrimination. A causal influence is discriminate if it systematically strikes one type of individual more often than it does other types of individual. Otherwise, it is indiscriminate.

To qualify as a discriminate causal influence over population members' reproduction, a causal influence must be systematically associated with some type of individual in the population. The systematic association must arise from some structural feature of the population, such as linkage between alleles, or a systematic association between types and an ecological causal influence produced, say, by homing tendencies among population members. What is critical is that should a causal influence simply turn out to differentially affect the reproduction of types in the population in some generation, it should not be counted as a discriminate causal influence. Discriminate causal influences are causal influences that are statistically associated with types in the model because of structural features of the population, such as linkage or habitat imprinting.

Pervasive causal influences, because they affect all population members, automatically count as indiscriminate in their influence. They cannot affect different population members at different rates because they affect all population members. It is important, too, to keep the notions of discrimination and interactivity apart. A discriminate causal influence need not have a different influence on different types of individual; it must merely be strike one type more often than others. Interactive causal influences need not be discriminate either; they can strike different types at the same rates while having different influences on different individuals.

5.2.3 Key concepts of the second categorization scheme

The three criteria of interactivity, pervasiveness, and discrimination yield a breakdown of six different sorts of causal influences. The figure is six and not eight, as one might initially suppose, because there can be no discriminate pervasive causes. Of these six initial sorts of causal influences, two play no role in the cause-to-model algorithm for generating *deterministic* models. Pervasive non-interactive causal influences do not have any direct impact on the dynamics of populations of competitors. They affect different types of individuals in the same way, and they affect all individuals. If they exist within a population, they need not be wholly without an impact on population dynamics, however. They could matter to census population size by serving to regulate the number of offspring that all population members have, and since census population size is nearly always relevant to effective population size, the "drift" variable in population genetics, pervasive non-interactive causal influences can impact the extent to which population dynamics are a stochastic affair. But I set these causal influences aside for now as irrelevant to the matter of laying out the cause-to-model algorithm. In chapter 7, I discuss how census population size impacts effective population size.

Another sort of cause that is not featured in the cause-to-model algorithm for generating deterministic classical models is that of non-pervasive, non-interactive, indiscriminate causal influences (NINPICs). NINPICs are responsible for introducing a stochastic element into population genetics models. It is because populations are beset by NINPICs that it makes sense to model their dynamics as a matter of chance. The impact of NINPICs is quantified using effective population size parameters; in the most interesting case of non-neutral alleles, the *variance effective population size* parameter is used to model the impact of NINPICs.

The cause-to-model algorithm initially generates "deterministic" models, one in which the impact of NINPICs has been idealized away. In deterministic models, NINPICs are treated in each generation as having their most likely impact on population dynamics, which is no impact whatsoever. When it is possible to do so, converting a deterministic Wright-Fisher to a stochastic model requires going through the entire process of generating the deterministic model and then adding the stochastic element into the picture at the very end, once the deterministic model has been fully developed. I will discuss how this is done in chapter 7 after the algorithm for deterministic equations is presented. For now, what is important is that NINPICs have no impact over how the deterministic portion algorithm works; even if one wants to incorporate the impact of NINPICs on population dynamics, one must generate a deterministic model first, and then add an additional step to include the impact of NINPICs through the introduction of stochasticity.

That leaves us with four categories of causal influences that must be countenanced in the deterministic portion of the cause-to-model algorithm. These are pervasive interactive causal influences, non-pervasive interactive discriminate causal influences, non-pervasive interactive indiscriminate causal influences, and non-pervasive non-interactive discriminate causal influences. These last four categories of causal influence are not yet enough, however, to categorize causal influences sufficiently narrowly for use in the cause-to-model algorithm. They must be broken down further in terms of their source.

In classical population genetics, causal influences over population dynamics, ones that may fall into any of the above classifications, have four main sources. Causal influences arise among alleles within the same genome; these lead to *gene-by-gene causal influences*. Causal influences also arise from individuals in the population causally influencing one another, as, for

instance, altruists do; I call these *individual causal influences*. Causal influences may arise from the ecological environment; and I call these *ecological causal influences*[21].

One last causal influence over population dynamics that requires separate consideration is sex; population geneticists must countenance s*exual causal influences*. Population genetics textbooks nearly always feature a discussion of sex-dependent selection, and usually present sex-dependent selection as occurring when the relative fitness of at least one genotype is different in males and females (e.g., Ewens 2004, 45). By the inversion principle, sex-dependent variation in the causal influence of genotypes signals that sex differences themselves are a causal influence on the development of zygotes.

In all, that means we will have to consider gene-by-gene causal influences, individual causal influences, ecological causal influences and sexual causal influences, where many of these may count as any one of the four basic types of causal influences picked out according to the interactive/pervasive/discriminate scheme. The algorithm handles gene-by-gene, individual, ecological, and sexual causal influences in very different ways, but gene-by-gene causal influences are especially unusual, so I discuss them first.

Gene-by-gene causal influences are only possible because genes are found together within gametes, zygotes and mating pairs. So I will begin with a discussion of the individuals that appear in the diploid lifecycle since an understanding of gene-by-gene causal influence hinges on this. That discussion will yield definitions of the different sorts of individuals in the diploid lifecycle. Defining the different sorts of individuals one finds in classical population

[21] I say "ecological" rather than "environmental" because, in one sense of "environment," the environment of an allele includes more than just its ecological environment, and may include its genetic context too. Some writers have claimed that gene selectionism involves treating alleles at other loci as part of the environment of some allele of interest (e.g., Kitcher and Sterelny 1988). The sense of "environment" in which alleles at other loci and ecological factors are both part of the environment of an allele is not a sense of environment that we need in the cause-to-model algorithm.

genetics causal terms will enable us to state the algorithm that follows by making reference to them.

5.2.3.1 Lifecycles and aspects of fitness. One of the crucial differences among classical population genetics models lies in the sorts of relative frequency terms they feature. Organisms with different sorts of lifecycles are tackled through different sorts of models featuring different relative frequency terms. The commitment to causal interpretation undertaken in this work extends to understanding relative frequency variables in causal terms.

My approach to relative frequency variables in classical population genetics is to treat their deployment as licensed when two or more competitors are causally related to one another in a specific fashion. I use causal language to say how competitors must be related so as to form gametes, zygotes and mating pairs, or better, how they must be related to form the *sorts* of composite bodies of which gametes, zygotes and mating pairs are our paradigm instances. I undertake to specify this relationship now because we must do so before we can consider how such things as ecological and individual causal influences must be modeled.

In classical population genetics modeling, gametes, zygotes, and mating pairs are treated as distinct types of entities picked out by distinct relative frequency terms. These individuals are arranged in terms of a lifecycle: among some organisms, haploid gametes pair to produce diploid zygotes, and among some of those organisms, the zygotes pair up to produce mating pairs.[22] The lifecycle is usually discussed in reference to the different *aspects of fitness*. Gametic, zygotic, and fertility selection are three of the four traditional aspects of fitness (Christiansen and Prout 2000).[23] The distinction between these aspects of fitness is parasitic on the distinction between

[22] Gametes fuse to form zygotes among polyploids, too; however I will be concerned specifically with the diploid lifecycle, as this is what population geneticists typically consider.
[23] The last "aspect of fitness" is sexual selection. Sexual selection is an aspect of selection in the same sense as gametic, zygotic, and fertility selection are aspects of selection, but the reason why this is so will become clear later.

the types of individual whose relative frequencies the fitness coefficients weight: the notion of gametic selection is parasitic on that of gamete, zygotic selection is parasitic on "zygote," and fertility selection is parasitic on "mating pair." So an understanding of how these entities are related in the lifecycle is critical to population genetics modeling because these different entities are all, at least potentially, weighted by different relative fitness coefficients (as well as other parameters).

The existence of lifecycles among population members has the further consequence that one cannot claim that gametic, zygotic, and fertility selection parameters reflect the causal influence on the *reproduction* of the entities with which they are associated. Only in the context of very simple models, ones featuring only zygotic selection, could we say that zygotic fitness parameters represent the causal influences of the zygotes on their reproduction, provided one means *reproduction by* the zygotes rather than *reproduction of* the zygotes. But we cannot say the same thing about relative fitness coefficients that are paired with gametes; among diploids, gametes paired with higher relative fitness values are more likely to form zygotes than are their less favored rivals, but they are not necessarily more likely to produce descendants, because they could tend to form rather unfit zygotes. Indeed, a zygote may produce gametes of a type from which it was not formed, as happens if it is a double heterozygote in which recombination occurs. So it does not even generally make sense to think of gametes as producing descendants *via* zygotes; the gametes that make up a zygote need not produce descendants at all and so we should not think of classical population genetics models as ones that feature descendant-producing gametes.

For now, I will simply note that there is no distinct type of entity with which assortative mating parameters, the ones that reflect "sexual selection," are paired. These parameters weight zygote frequencies, but do so in a different way than do zygotic selection coefficients; one difference between zygotic selection coefficients and assortative mating parameters is that only the latter, weighted by the relative frequencies of the mates, must sum to one.

Similarly, in deterministic models featuring variables for mating pair frequencies, zygotes with higher fitness values form more mating pairs, but they do not necessarily reproduce more because they may form especially infertile mating pairs. So we cannot make sense of the fitness coefficients with which zygote frequencies are paired as ones that quantify the causal influence of the zygotes on *reproduction*. Furthermore, just like gametes, zygotes that form mating pairs may produce descendants of a different type than either zygote that formed the mating pair: Two homozygotes for different alleles will produce only heterozygote offspring for instance. So we cannot say that relative fitness values represent the causal influence of zygotes over zygote reproduction, where we mean the production of zygotes of the same type.

What I will say instead is that fitness coefficients quantify the causal influence of individuals over their *progress on to the next stage of the lifecycle*, or more simply their progress. For gametes, progress consists in zygote formation. For zygotes, progress consists in mating pair formation (in some models) or gamete production (in other models). For mating pairs, progress consists in zygote production. Some of these stages of the lifecycle can sometimes be ignored as irrelevant to population dynamics, and all of the stages do not occur among at least some organisms, but all the stages must be recognized as distinct in an account of selection theory that is broad enough to cover standard population genetics models featuring diploids.

I should note, too, that while Christiansen and Prout divide up the components of fitness into multiple aspects, a move that is crucial for understanding the deployment of population genetics models, those authors make the mistake of defining these components of fitness in terms of the *outcomes* of selection processes: zygotic selection is the differential survival of zygotes to maturity, gametic selection the differential survival of gametes, and so on (2000, 148). It is of prime concern in this work that the concepts at work in selection theory be construed such that

they can function as part of explanations of population dynamics, and the outcome-based definitions offered by Christiansen and Prout undermine the explanatory power of equations in which variables representing the various components of selection appear. We do not want our, say, zygotic fitness variables to refer to the actual survival rates to maturity of zygotes, because we would then have to have a grip on what happened in the zygotic phase of the lifecycle before being able to determine zygotic selection coefficients for population members. The same goes for the rest of the stages of the lifecycle. If we want to explain populations' dynamics, we must be in a position to infer it in a non-circular fashion, and the definitions offered by Christiansen and Prout stymie this ambition.

5.2.3.2 MICERs. We need to know what gametes, zygotes, and mating pairs are before we can set out the algorithm for generating population genetics models. Indeed, we will want a grip on what *sorts of things* these individuals are, because the algorithm is supposed to be a cause-to-model algorithm, one that takes makes use of causal information as an input. This means we will need to characterize what gametes, zygotes, and mating pairs are in *causal terms*. In a sense that will be specified in this section, gametes, zygotes and mating pairs are the causal products of a specific kind of causal relationship between competitors. We can get a grip on when it is appropriate to deploy a model featuring gametes, zygotes, and mating pairs by getting a grip on the special way that competitors are causally related when they together make up a gamete, zygote, or mating pair.

That it is a specific sort of causal relationship that allows us to recognize whether a system features individuals of specific kinds might seem like a rather bizarre contention. Surely we do not recognize that a system features, say, mating pairs on the basis of a peculiar sort of causal relationship that exists among the alleles carried by the pairing organisms. The same goes

for gametes and zygotes. I can tell that something is a zygote without checking how the alleles it bears are related. I am sure that everyone reading this is a zygote, for instance. However, what I seek in this work are explanatory definitions; I seek definitions that pick out entities in terms of the features they have that lead them to play the roles in the theory that they do. Gametes are not treated in selection theory in the way that they are treated because they are gametes, they are treated in the way that they because of how their component alleles are causally related.

Consider a parallel: while it is the case that all known rights-deserving entities, people, are also humans, it is not the case that people deserve rights because they belong to a particular clade in the latest taxonomic hierarchy developed by professional systematists. Indeed, it is not inconceivable that we might find, elsewhere in the universe, entities that deserve rights despite *not* being humans. Evolutionary history could have gone slightly differently, too, such that other lineages of *homos* survived into the present era, and could talk and play the rights-and-responsibilities the rest of us do. Their status as members of a different clade would hardly legitimate their enslavement. So it is surely at least an evolutionary accident, if not a cosmological one, that all rights-deserving entities turn out to be *Homo sapiens*.

An account of why rights-deserving entities are such should not make appeal to their status as humans, even were the case that all and only humans deserve rights, something that would make it possible to assess rights-deservingness on grounds other than personhood. Similarly, even though we can spot zygotes without doing so on the basis of a peculiar causal relationship into which the alleles they bear have entered, it is *because* of that peculiar relationship among their constituent competitors that zygotes play the role they do in selection theory. We need not invoke an explanatory definition in order to deploy a notion every time we deploy it. But we can still recognize that the definition supplies the *justification* for a notion

playing a specific role in reasoning, even though we can deploy the notion without using the definition.

Alleles that are bound together in gametes, zygotes, and mating pairs are engaged in a special, and temporary, causal relationship with two components. The first component is *mutual investment*. Each stage of the lifecycle terminates in a causal bottleneck: there is no way for an allele carried by, say, a dog zygote to produce descendant alleles except insofar as the dog zygote bearing it finds a mate. That mating pair, too, had better be fertile for the alleles in both the male and the female to have any hope of producing descendants. I dub this aspect of the relationship between alleles within the same individuals an instance of *mutual investment*. The reproduction of all the alleles, or more generally all the competitors, within a single individual is contingent upon the progress of the individual to the next lifecycle stage. Alleles that are related such that their descendant production is each contingent upon the same event exhibit mutual investment.

Alleles forming parts of the same individual have a further feature. Whatever competitive causal influences that competitors within an individual have one another, these causes have *no impact* on the progress of the individual bearing them. Alleles within an individual cannot engage in activities that *both* 1) causally impact the event upon which their descendant production is both contingent and 2) function to increase their likelihood of producing descendants while inhibiting others that share the same causal bottleneck. Note that I am not saying that competitors within an individual cannot both engage in competition and impact the reproduction of the individual bearing them. I claim that they cannot do both these things *by means of the same causal mechanism*.

This last feature of how alleles within an individual are related is a way that members of subgroups of the sort featured in "group selection" models are *not* related. Some subgroups may

be formed from individuals in such a way that the alleles in the subgroups are mutually invested in the fate of the subgroup as a whole. Insofar as that is the case, subgroups and individuals are similar. However, members of the same subgroup can, and indeed always do, compete with one another in a fashion that matters to the probability that any members of the subgroup will produce descendants.

For instance, an allele coding for virulence in the *mixoma* virus will produce more descendants within a single infected rabbit than will an allele coding for avirulence (this example is drawn from Sober and Wilson 1998). This behavior is competitive: virulent strains of *mixoma* produce descendants at the expense of avirulent ones by using up rabbit resources to do so. But this behavior also has an impact on whether any of the viruses in the host rabbit go on to produce any future descendants at all. The *mixoma* virus is passed on by mosquitoes that only bite live rabbits. By producing more descendants than do rival strains, virulent strains kill the infected rabbit sooner and hence decrease the probability that the rabbit will be bitten by a mosquito, something that must occur for *any* strain within the rabbit to be carried to other rabbits where viral replication can continue. The virulence of the strains in a rabbit determines *both* whether any of them make it outside of the context of the subgroup and what proportion of those that do are virulent and avirulent.

In contrast, the causal chain by which the alleles in a zygote causally influence the progress of the organism is an autonomous one, one which has nothing to do with how many descendants the alleles produce compared to other alleles within the zygote. The developmental impact of an allele, and hence its impact on the progress of individual of which it is a part, is independent of whatever else the allele might do to generate more descendants than does a rival alleles, such as distorting meiosis through meiotic drive. The descendant production of alleles in

an individual is contingent upon the progress of the individual of which they form a part, and the alleles cannot do anything, that is, they cannot do *any one thing*, that both 1) influences the progress of the individual that bears them and 2) influences their rate of descendant production relative to other alleles within the individual. Two alleles within a single zygote, then, are such that a single event (the reproduction of the zygote) is a necessary for both to produce any descendants and further are such that they cannot causally influence that event by means of some activity that inhibits the descendant production of the other allele. The same goes for alleles within the same gamete or mating pair.

I call the relationship between competitors forming an individual a MICER, a mutually invested, competition-excluding relationship. When relationships of this sort occur, as they do among alleles that together form parts of the same gametes, zygotes, and mating pairs, they mandate the deployment of relative frequency variables that refer to individuals formed by MICER related alleles and distinguished by the competitors they bear. So, for instance, we can tell whether it would make sense to deploy a model featuring zygote frequencies by asking whether any competitors within haploid gametes form a MICER with each other. Equally, we can tell whether it would make sense to deploy a model featuring mating pair frequencies if the zygotes in the population engage in a MICER. Even gametes are formed from MICERs, in this case, relationships of mutual investment and competition exclusion among alleles at different loci.

5.2.3.3 Aside about meiotic drive. That alleles within a zygote are related by MICERs is not immediately obvious because rival alleles from a single heterozygote clearly compete with one another when one of the alleles is a driving allele, and MICERs are supposed to be competition-

excluding relationships. This fact, however, does not endanger the stance that alleles at the same locus in a heterozygote still form a MICER.

Technically, meiotic drive is a case of gametic selection. An allele may bias the rate at which it forms zygotes by biasing meiosis as in meiotic drive or, more commonly, by killing rival gametes, as do killer sperm. But the fact that alleles can do this can be accommodated on the view undertaken here. The relationship of two alleles that invest in a zygote is always temporary, and the relationship ends as the next stage of the lifecycle begins. Meiosis marks the beginning of the gametic lifecycle stage and the causal influence of an allele over meiosis is an instance of gametic selection, not zygotic selection.

That last claim is not an idle one. I am not carving up the lifecycle arbitrarily just so that I can maintain that meiotic drive is an instance of gametic selection. For one thing, any textbook will present it as such. But for another, it matters when modeling meiotic drive mathematically that it counts as gametic selection. Gametic selection coefficients weight the frequency terms for gametes produced by heterozygotes. Gametic selection coefficients do *not* weight the contributions that heterozygotes make to the average fitness of zygotes. That means that the relative fitness coefficients used to represent drive do not quantify the causal influence of the zygotes on their relative reproduction rate. Godfrey-Smith and Kerr (2002) overlook this, and this error is one source of their mistaken claim that genotypic selection is an instance of group selection on alleles.

The fact that heterozygotes have relative fitness parameters that do not reflect whatever meiotic drive may occur between the alleles they bear brings out the special role of individuals in selection theory. Alleles in a zygote, even driving alleles, have an impact on the development of zygotes that is distinct from whatever influence they may have over meiosis or fertilization.

MICERs are formed out of causal relationships; the contributions made by alleles that are part of the same genotype to zygote development are not causal influences through which they can compete with one another. Whether alleles function as causes in *other* ways, beyond contributing to zygote development, is not relevant to whether they engage in a MICER. So alleles may case the deaths of rival sperm, and more generally can influence the gamete formation process such that they form parts of more gametes than do other alleles competing for the same locus. They can do these things *in addition* to forming a MICER with those other alleles. The processes of zygotic selection on genotypes and gametic selection on gametes are independent.

5.2.3.4 Distinguishing haploids, gametes, zygotes, and mating pairs. So far, I have argued that the special gene-by-gene causal influences that occur among competitors within gametes, zygotes, and mating pairs are causal relationships of a particular sort, mutually investing, competition-excluding relationships. This relationship is defined causally, so it is a generalization of the sort of relationship that alleles have to each other when they form part of the same individual; things other than alleles, anyway, could instantiate the relationship. But all we know now is that competitors that together make up gametes, zygotes, and mating pairs always form MICERs; I must still distinguish gametes, zygotes and mating pairs among the individuals whose competitors are related in this way. Once this is done, we will have notions stated in causal terms that generalize "gamete," "zygote," and "mating pair," and we will be in a position to use these notions in the cause-to-model algorithm.

The strategy pursued here will be to first get a grip on the notion of gamete and to use this to get a grip on the notions of zygote and mating pair. Diploid zygotes are created by MICERs formed from competitors found in two gametes, while mating pairs are created by

MICERs formed from competitors found in two zygotes. So the critical question is, what are gametes?

To get a grip on the notion of gamete, we must recognize that the alleles within a gamete are not members of the same population. One of the features of the definition of population offered earlier in chapter 4 is that only types of entities that are *competing* count as members of the same populations, and alleles at different loci cannot compete. That alleles at different loci are actually found in different populations has the clear, though not trivial, implication that their relative frequencies need not add to 1. But the fact that alleles at different loci are not part of the same population also has importance for discerning what gametes are.

Distinct alleles within the same gamete may be related by MICERs even though they are not members of the same population. Zygotes and mating pairs always include competitors that are members of the same populations, alleles that are competing for the same genetic loci. So we can pick out gametes as those peculiar individuals that bear competitors that are 1) related by MICERs, and 2) are not members of the same populations.

Using the notion of a MICER, along with the notion of population defined earlier we can pick out gametes in causal terms as the product of competitors from different populations that are related by MICERs. Zygotes, then, are MICERs formed from pairs of MICERs formed from competitors from different populations, while mating pairs are MICERs formed from pairs of those.

Any competitors that do not form any MICERs at all, with members of the same or different populations, will be treated as haploids are treated in population genetics models. MICERs are temporary relationships, and multiple genetic variations in haploids are *permanently* engaged in relationships of competition exclusion and mutual investment. Not only are alleles in a haploid individual mutually invested in the reproduction of the larger individual, all their descendants will equally be mutually invested in the reproduction of the same individuals too. In contrast, the causal relationships between different alleles at different loci in diploid individuals are impermanent; they eventually break down through recombination. Indeed, that this is what triggers the deployment of multi-locus models is clear from the fact that two-locus models reduce to multi-allelic models as recombination rates approach negligible rates (Hedrick 2005, 561). Since the lineages of two alleles in the same haploid genome are inexorably bound together, and the relationship of competition-exclusion and mutual investment that is used to define gametes, zygotes, and mating pairs is, by definition, a temporary relationship, alleles within haploid do not form MICERs.

That the deployment of models featuring haploids is triggered by the existence of competitors that do not form MICERs helps explain the scope of such models. They work just as well for competition between lineages of organisms from different species that do not interbreed as they do for lineages of variant alleles at a locus in haploid organisms that do not engage in sexual reproduction. Indeed, that we can be indifferent to what we count as the competitors when dealing with haploids—is the competitor the organism or some subset of its genes?—is equally explained by the permanence with which genetic variations in the same haploid individual are stuck together. It does not matter whether we consider the organisms or its peculiar genetic variations as the rival competitors when dealing with haploids because the lineages of the genetic

variations and the organisms necessarily overlap. It is really the models that involve sexual reproduction that are special, and they are special because they involve competitors that form MICERs. The "haploid" models are suitable for all the other sorts of competitors.

5.2.3.5 Generalizing the second categorization scheme. Earlier I picked out gene-by-gene causal influence as one of four sorts of causal influence that must be addressed in the cause-to-model algorithm. It turned out that gene-by-gene causal influences were a specific sort of causal relationship that competitors could have to one another, relationships of mutual investment and competition exclusion. Causal influences emanating from alleles and acting on other alleles need not take the determinate form of MICERs, however. Just because something is an allele does not mean that its causal connection to any other allele can be characterized as one of mutual investment and competition-exclusion. In short, alleles can exert causal influences on alleles in other individuals, too. Frequency-dependent selection models are deployed to handle causal influences of this last sort. In classical population genetics, we weight relative frequency terms for individuals by relative fitness functions whose arguments consist in weighted relative frequency terms for individuals at the same point in the lifecycle. Those functions must be interpreted causally. So if I set the relative fitness of a homozygote according to the following function (as in Hedrick 2005, 223):

$$w_{11} = p^2 w_{11\cdot 11} + pq w_{11\cdot 12} + q^2 w_{11\cdot 22}$$

Equation 5.9

each of the RHS terms picks out the causal influences of each type of zygote on the progress of the A_1A_1 homozygote.

That frequency-dependent selection models are causally interpretable is important because, technically, the causal influences we model using frequency-dependent selection

models emanate from genetic variations; the arguments of the functions ultimately quantify the causal influences of alleles over other alleles. But to keep them distinct in the algorithm that follows, I call these causal influence *individual causal influences*. I do this just to keep things straight, even though such causal influences have their ultimate source in variant competitors.

This means that out of the four types of causal influence in the second categorization scheme, two have so far been understood entirely in terms of the base language of this work. What are known as gene-by-gene causal influences are a specific type of causal influence among competitors, a MICER. Other causal influences among competitors are what we will call individual causal influences. That leaves sexual causal influences and ecological causal influences. Can "sex" and "the environment" be understood in causal terms? In fact, we need not understand both in causal terms. All we will need to do is understand one of these in causal terms, and the other can be understood as the correlative to the other three types of cause in the second categorization scheme. We will thereby have articulated the second categorization scheme in entirely causal terms.

Sexual causal influences can be understood without reference to the concept of sex. Sex differences are manifested in population genetics models as restrictions on MICER formation. Quite simply, competitors in one sex cannot form MICERs with competitors that are found in individuals of the same sex.[24] So we can generalize sex differences as barriers to MICER formation. We can say what it means to be of a sex *X* for an arbitrary *X* by appealing to what it means to form a MICER, something which has been presented in causal terms: Individuals of sex-value *X* cannot form causal relationships of mutual investment and competition-exclusion with individuals of sex-value *X*.

[24] Gametes have sex-of-origin differences that are equivalent to sex differences in the fashion being specified here.

Having understood sex in the fashion just rehearsed, we can understand "ecological" causal influences as all the causal influences modeled in selection theory except for MICERs (gene-by-gene causal influences), individual causal influences (gene-by-gene causal influences that are not MICERs), sexual causal influences (causal influences due to differences manifested by barriers to MICER with other population members), and ecological causal influences (everything else). We now have our second categorization scheme laid out in the base vocabulary of this work.

5.2.3.6 Substructures and subgroups. Population geneticists consider a variety of different sorts of groupings of competitors and individuals. The most significant sorts of grouping they consider are formed from competitors engaged in MICERs, mutual investing, competition-excluding relationships. Competitors that are part of the same gametes, zygotes, and mating pairs are so related. The formation and dissolution of MICERs is represented explicitly in the lifecycle graphs. But competitors and other individuals form other sorts of groupings, too. The two main sorts of groupings we need to consider henceforth in the algorithm are substructures and subgroups.

Substructures are permanent features of the landscape inhabited by populations. Population members in different substructures migrate between substructures at fixed rates; this is why members of different substructures still count as members of a single population. But, except for the descendants of the migrants, the descendants of members of one substructure appear in the same substructure as that in which their parents initially appeared. Crucially, barriers between substructures act as barriers to MICER formation: the gametes in one substructure form zygotes only with the gametes in the same substructure, and the zygotes in one substructure form mating pairs only with their fellow substructure members, too.

The barriers between substructures may form barriers to other causal influences too; members of different substructures may not directly compete with one another.[25] This will impact the relative fitness parameter deployed in the substitution phase of the algorithm, and I discuss this at length in the substitution phase of the algorithm.

Insofar as members of a substructure form MICERs only with each other, our rules for drawing lifecycle graphs will require that gametes and zygotes share a substructure when forming zygotes and mating pairs, respectively. Similarly, later on, when we come to consider substitution rules that countenance ecological, sexual, and individual causal influences, we will evaluate the extent to which these causal influences are pervasive and the extent to which they are discriminate with respect to specific substructures. That substructures play these sorts of roles in the algorithm makes them different from *subgroups*.

Subgroups are temporary groupings of individuals, ones that last no longer than the gametic stage of the lifecycle or the zygotic stage of lifecycle (in principle, mating pairs could form subgroups too, but one never sees this). Individuals that end up in one subgroup may form MICERs with individuals from other subgroups; this is what makes subgroups different from substructures, they dissolve before the next lifecycle stage begins. Subgroups may be familiar to philosophers from so-called group selection models (e.g., Wilson 1990; Sober and Wilson 1998; Godfrey Smith and Kerr 2002). Subgroups are handled in the substitution stage of the algorithm and are not recognized when lifecycle graphs are drawn.

Some temporary conglomerations of individuals exist only as a means to classify together individuals subject to the same causal influences. To keep things straight, I will call these last

[25] Recall from chapter 4 that the descendants of co-temporaneous individuals that do not compete directly may compete, and so, by the transitivity of competition, cotemporaneous individuals in different groups may compete indirectly with each other despite failing to compete directly.

sorts of things *partitions*. Partitions are formed from individuals that share nothing more than subjection to the same non-pervasive ecological causal influences. Partitions of these sorts are mere mathematical classificatory tools put in place to accurately model ecological causal influences; they are not subgroups and we will not be concerned with them until section 6.3.5.2.

In contrast, true subgroups exist as independent groupings, ones that are not simply formed out of whichever individuals happen to be subject to the same ecological causes. These subgroups are created by barriers to causal influence between members of different subgroups. Such barriers have at least one of two consequences. At a minimum, subgroups provide barriers to individual causal influences. Some population genetics models feature individuals that causally influence other cotemporaneous individuals in the system; "group selection" models featuring altruistic and selfish types are prominent among these. However, because subgroups form barriers to causal interaction, they may form barriers to *competition* too, since competition is just a special kind of causal interaction. Such barriers to competition have implications for the average fitness parameter that is appropriate for a model featuring subgroups, as I discuss in the next section.

5.3 LIMITS TO THE ALGORITHM

I have set out the key concepts we will use to breakdown all the different causal influences we consider in the cause-to-model algorithm. We have a grip on the four fundamental sorts of causal influences that matter to the cause-to-model algorithm no matter what their source: pervasive interactive causal influences, non-pervasive discriminate non-interactive causal influences, non-pervasive discriminate interactive causal influences, and non-pervasive, indiscriminate interactive causal influences. We have further distinguished as gene-by-gene causal influences,

individual causal influences, sexual causal influences and ecological causal influences. We have also used the official causal language of this work to get a grip on what gametes, zygotes, and matings pairs are. Having a grip on our main types of causal influences *and* having a grip on the entities that will be represented as subject to them, we understand the concepts we need to write down the algorithm for generating population genetics models on the basis of causal information about populations.

Before getting to the algorithm itself, I want to set out more fully the limitations of the algorithm. The algorithm is not fully general, it does not yield an equation for any sort of population whatsoever, and I want to at least mention up front what sorts of populations it can handle and what sorts it cannot. First of all, the algorithm works for discrete generation, "Wight-Fisher" models. Some of these, those that are recursions on two types of gametes and which do not involve individual causal influences, can be treated using equations from diffusion theory, in which evolution is approximated as a continuous process by imagining ever shorter generations (Rice 2004, ch. 5). But the algorithm discussed here does not yield Moran models, Cannings models, or age-structured models. Equally, it does not yield models that deploy the formalism of quantitative genetics to track population dynamics, what Lewontin calls the "biometric approach" (1974, ch. 1). While these limitations are certainly limitations, it is worth noting that, even in advanced population genetics textbooks such as Ewens (2004), the vast majority of models of non-neutral population dynamics are Wright-Fisher models. Wright-Fisher models have been the focus of population researchers' development of quantitative models of causally complex situations, especially models that serve to demystify adaptation, altruism, and genetic polymorphism.

I also note that Wright-Fisher models are the most flexible population genetics models when it comes to representing the impact of contextual causes on population dynamics. Cannings models work only for neutral evolution; Moran models work only for haploids, the biometric approach generates conclusions about long-term population dynamics only for populations with an unlimited supply of genetic variations contributing to the trait under selection, which is an unrealistic assumption. Wright-Fisher models are actually some of the hardest models for which to state a cause-to-model algorithm, because they can handle the greatest breadth of causally diverse populations. This fact is equally linked to their traditional role in demystification.

Another limitation to the algorithm is that I consider models of haploid, haplo-diploids, and diploid populations only; I do not consider populations that exhibit polyploidy. This is once again largely because of the traditional focus of the population genetics research community on haploids and diploids. I simply could not find out how selection on polyploids worked from my population genetics textbooks, as well as a little research into the primary literature on polyploidy. Much the same goes for bacterial populations that do not mate as some diploids do, but exchange alleles through any of a number of "parasexual" processes that I have not yet learned how to model. My hope is that coming to terms with these unusual populations will amount to understanding how to draw appropriate lifecycle graphs for them, such that integrating models of selection in these unusual cases will not require adjustments to the substitution rules that come later.

The algorithm is further limited insofar as I use fixed parameters for quantities that population geneticists have sometimes treated as functions. The algorithm does not countenance homing parameters that are functions of relative frequency terms, for instance, though these are widely considered in models of spatially variable selection (e.g., Hedrick 1993). Equally, I use

unspecified functions to weight relative frequency variables in the face of individual causal influences, even though such functions must be specified for inferences about population dynamics to be made. The algorithm is not incompatible with the use of functions for these parameters, but it does not mandate them.

Another serious assumption I make is that individual, sexual, and ecological causal influences do not interact with each other, though they may interact with the type differences of the individuals they strike. This assumption is an extension of the textbook assumption of the multiplicative collapse of viability (Hedrick 2005, 176); I dub my extended version of this assumption the *multiplicative collapse of fitness assumption*.

One last feature of the algorithm is worth mentioning here. In stating the algorithm, I resort to the traditional language of population genetics. Even though I went out of my way to define some crucial bits of causal and statistical terminology, and even though I bothered to show how one can manufacture generalizations of the notions of gamete, zygote, and mating pair from causal notions, and even though I bothered to re-interpret gene-by-gene, individual, sexual, and ecological differences in entirely causal terms, I will nevertheless put forward the algorithm using the more determinate biological versions of these concepts. So I will talk about alleles instead of competitors, gametes instead of individuals formed from competitors in different populations involved in a mutually invested, competition-excluding causal relationship, and so on.

While my adoption of biological vocabulary in stating the algorithm clearly limits its scope in a fashion that is at odds with the intent of this work to *generalize* selection theory using cause-talk, at least it will be convenient to talk about mating pairs instead of MICERs formed from MICERs formed from competitors in distinct populations. Furthermore, the algorithm can

be generalized simply by substituting the official understandings of its various bits of terminology for the traditional vocabulary actually used in the presentation. Finally, I composed the algorithm in the more determinate language of population genetics textbooks because if I ever make use of the algorithm outside the context of this dissertation, it would have to be versed in the traditional language of population genetics.

Intellectual honesty provokes me to mention all the above limitations to the cause-to-model algorithm before presenting it. But it is easier to discuss how severe these limitations are after the algorithm has been presented, since it will be clearer then how these limitations function as such. So I postpone further consideration of them until after the presentation of the algorithm itself.

6.0 THE CAUSE TO MODEL ALGORITHM

The algorithm I offer for generating classical population genetics models is a multi-step affair. Its main steps are these: 1) the deployment of a decision-tree to determine which among several fundamentally different types of classical population genetics model is appropriate for some system; 2) the drawing of some specialized directed acyclic causal graphs, ones that portray the lifecycle of the target systems and from which equations can be derived that feature an appropriate number of relative frequency terms referring to appropriate types of individual; 3) the deployment of substitution rules upon the equations yielded by the graphs, substitutions that generalize the equations such that the remaining causal influences not tackled at earlier stages of the algorithm, specifically individual, ecological, and sexual causal influences, can be taken into account; 4) the collapse of two systems of equations to generate a single recursive system of equations.

6.1 THE DECISION TREE

The first stage of the algorithm consists in a decision tree (figure 6.6) that functions to distinguish six basically different kinds of classical population genetics models. In this section, I show how to make the distinctions made in the decision tree. The stage, the graph-drawing stage of the algorithm, will treat each type of model distinguished by the tree using differently.

6.1.1 Haploid vs. diploid models

The definition of MICERs, one of the critical concepts needed for selection theory discussed in the previous chapter, is the first concept to be put to use in the cause-to-model algorithm. Classical population genetics for populations in which MICERs are formed from members of the same population, ones featuring zygotes and perhaps also mating pairs, are different from ones designed for populations in which MICERs are not formed among members of the same population. I'll call the latter sorts of model "haploid models," though they work fine for cases in which individuals in different non-interbreeding species function as competitors. If our competitors do not form MICERs with members of the same population, then we must deploy a haploid model; otherwise we need to know further details about our population to fix on an appropriate sort of lifecycle graph.

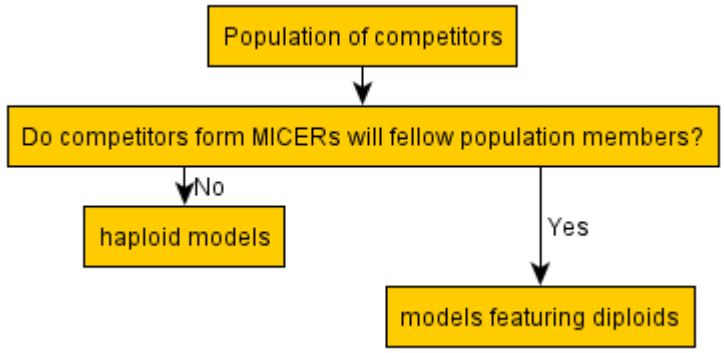


Figure 6.1 Haploid vs. diploid models

6.1.2 Fertility selection

The next thing we need to know in order to follow the decision-tree is whether our population forms mating pairs, and if so, whether the zygotes do so discriminately or interactively (or both). If the process of mating pair formation is either discriminate or interactive, such that different types of zygotes tend to form mating pairs with other zygotes in a genotype-sensitive fashion or such that zygotes do not make mate-independent contributions to fertility but instead have

different impacts on fertility depending upon the type zygote with which they mate, then we must deploy a model that is a recursion on zygote frequencies and features relative frequency terms for mating pairs; otherwise we can use a analytically more tractable model that is a recursion on gamete frequencies.[26] Different lifecycle graphs are used to generate equations that are recursions on zygote frequencies instead of recursions on gamete frequencies.

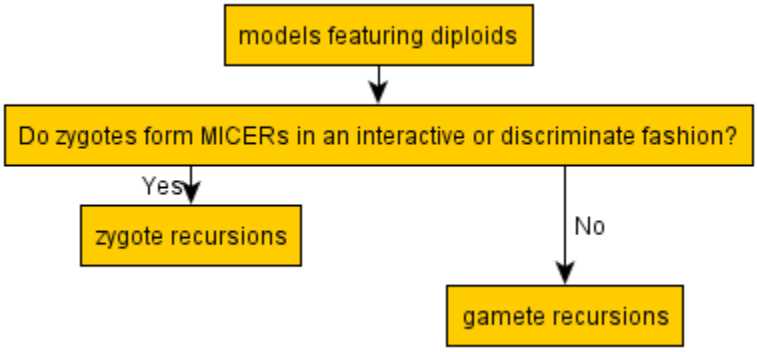


Figure 6.2 Fertility selection

6.1.3 Haplo-diploidy

The next thing we need to know is whether MICERs are formed from alleles in *both* sexes of a diploid species. *Hymenoptera* exhibit a peculiar genetic structure: the females are diploid while the males are haploid. That means the alleles among the males do not form MICERs with members of the same population, while those among the females do. Haplo-diploid models suitable for *Hymenoptera* are a unique sort of recursive model, and must be considered separately in the graph-drawing stage of algorithm. Haplo-diploid models are suitable for X-linked genes, too (Hedrick 2005, 75).

[26] Because introducing mating pairs into a model is done by assigning relative frequency terms to mating pairs, and it is these mating pairs of zygotes that then produce the zygotes of the next generation, models featuring mating pairs contain sufficient information to make it possible to connect this-generation mating pair frequencies with next-generation *zygote* frequencies. Mating pairs that are more or less successful produce more or fewer *zygotes*, not more or fewer *gametes*. That is why models that feature mating pairs are recursions on zygote frequencies.

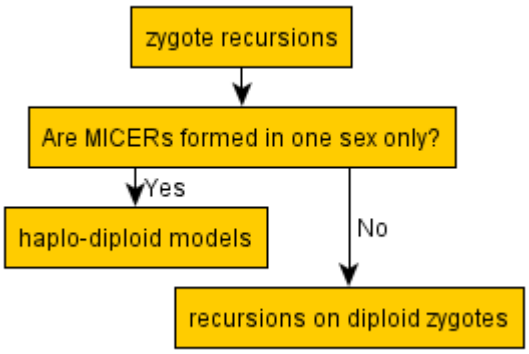


Figure 6.3 Haplo-diploidy

6.1.4 Sex-dependent selection

Next we must consider whether sex differences act as interactive causal influences on zygotes. If so, we must deploy a sex-dependent selection model so that it is possible to assign the males and the females of the different types in the model different relative fitness coefficients. Sex-dependent selection models are recursions on individuals in which zygotes are differentiated by their sex and gametes, if featured, are differentiated by their sex of origin. All diploid selection models come with sex-dependent and sex-independent versions, whether or not they are recursions on zygote frequencies or gamete frequencies. Note that it is not the case that sex-dependent models are appropriate for populations in which there are two sexes. Sex differences that have no interactive causal impact on the progress of individuals can be ignored.

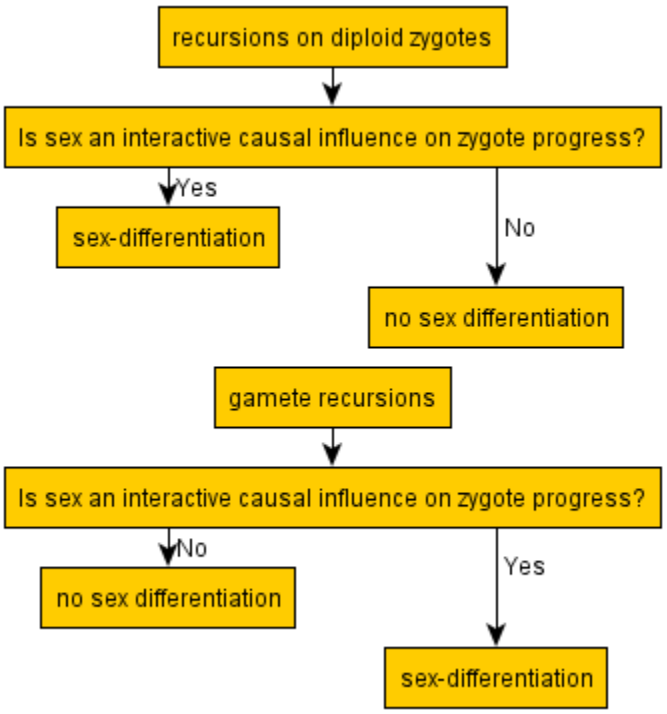


Figure 6.4 Sex-dependent selection

6.1.5 The Full Decision tree

This completes the first step of the algorithm. The decision tree just rehearsed terminates with 11 fundamentally different types of populations that are associated with 11 fundamentally different types of classical population genetics model. I note that there are more types of populations than this because I consider neither bacteria that engage in parasexual processes nor polyploids. The full decision tree is picture below:

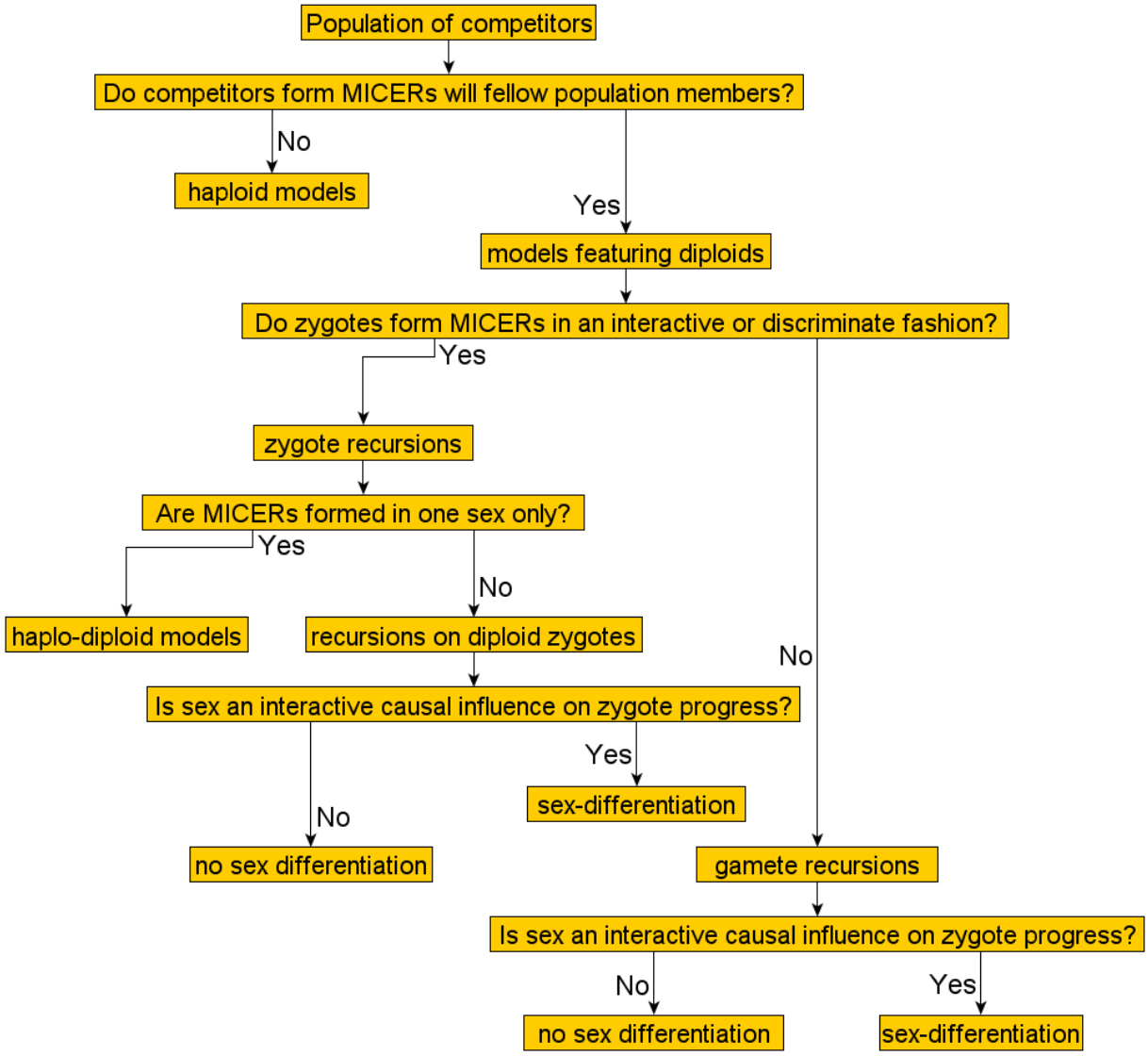


Figure 6.5 The full decision tree

6.2 LIFECYLE GRAPHS

The next step in the algorithm consists in instructions for drawing directed acyclic causal graphs to represent the lifecycle of populations. Each of fundamentally different type of population discerned in the last stage is handled differently in the graphing portion of the algorithm, so we will consider six sets of graphing rules for each of these different sorts of populations. Before actually giving the instructions for drawing graphs appropriate to the 11 different sorts of populations just picked out, I will discuss my use of causal graphs, the language in which they will be written, and what the graphing instructions allow one to do.

I avail myself of the sorts of graphs one finds used everywhere nowadays to formalize causal knowledge (Glymour, Scheines, and Spirtes 1993; Pearl 2000; Woodward 2003). I stress, however, that I do not represent the entirety of the causal information about populations in the lifecycle graph. Much of the causal information that will be fed into the algorithm is fed into the algorithm later on, specifically through the use of substitution rules at the next stage of the algorithm.

The main purpose of the graphing portion of the algorithm is to generate equations the right number of relative frequency terms for the right sorts of individuals in the model. The number of different types of individuals whose relative frequencies must be captured using variables in a model is a function of the number of different types of alleles in the population, as well as the number of loci used to differentiate the alleles, the number of substructures, and the presence of interactive sex differences. Distinct nodes in the graphs are assigned for each distinct type of individual, and edges are used to represent how they causally contribute to the individuals who make up the next lifecycle stage. Our models will reflect no more variant types of individuals than is necessary.

The other thing the graphs are used to do is to introduce some coefficients for relative frequency variables through weightings on edges that depict formation of the individuals in the next stage of the lifecycle. Meiotic drive parameters, recombination rates, and parameters representing discriminate MICER formation are all represented by weights on edges of lifecycle graphs. Once again, we add these only when necessary.

6.2.1 Inferring equations from lifecycle graphs

The purpose to which I put the directed acyclic graphs, specifically generating systems of equations, is an unusual use of such graphs. Directed acyclic causal graphs are not typically used to show the variables associated with each node are functionally related. As used by Pearl (2000), Glymour et al. (1993), and Woodward (2003), directed acyclic causal graphs display causal-functional dependencies of a mathematically indeterminate sort. Directed acyclic causal graphs are used to show that the value of one variable depends on the value of one or more other variables, but they include no information about the precise mathematical nature of this dependence. So a variable associated with a node in one of Pearl's graphs might take on a value that is the product of the values of the variables associated with its parent nodes, or perhaps their quotient, their sum, their difference, etc. A directed acyclic causal graph for some system merely shows that the descendant is *somehow* functionally dependent on its ancestors, but it says nothing about what sort of mathematical function relates them.

Things are different with the lifecycle graphs used here. These yield definite equations involving definite operators because I assign a definite rule that determines what sort of mathematical-functional dependency obtains between parents and children.[27] The rule has four components: 1) all coefficients on edges weight the variables associated with the nodes at their

[27] This is one of the main reasons why more causal information about a population is *not* represented in the lifecycle graph; adding more causal information would make it impossible to lay down general rules concerning how equations were to be inferred from the graphs.

feet by forming products with them; 2) when ancestors are composing a MICER, that is, when gametes are combining to form zygotes, or zygotes are combining to form mating pairs, values of variables for descendant nodes are the product of the values of their parents; 3) when a MICER is being broken down, that is, when zygotes are producing next generation gametes or when mating pairs are producing next generation zygotes, descendant values are the sum of parent values; 4) post-migration nodes and haploid nodes are always the sums of their parent nodes.

Lifecycle graphs are used to yield systems of equations, rather than a single equation. One equation is yielded for every leaf of the graph, that is, every node that has no descendants but does have ancestors. Furthermore, so that the substitution rules we deploy later on work correctly, each non-haploid population will be assigned two distinct lifecycle graphs, each graph representing a distinct stage of the lifecycle, such that two systems of equations are produced for the system.

6.2.2 Some conventions for lifecycle graphs

Stating rules for drawing lifecycle graphs can get complex, so here I rehearse a few graph-theoretic notions that will be useful in stating rules later. Many of the definitions below are standard ones; I propose a few of my own notions to use as shortcuts.

6.2.2.1 Heads, feet, paths, ancestors, children, parents, roots, leaves. The head of a directed edge is marked by an arrowhead; the foot has no arrowhead. Paths are sequences of one or more directed edges, whose component edges may point in any direction; I repeat, the edges that make up a path may point in *any direction*. An ancestor of a node is another node that can be connected to it by a path, all of whose component edges are pointed in the same direction, toward the descendant. Descendant is the reciprocal notion to that of ancestor. A node's parent is

its immediate ancestor; a node's child is its immediate descendant. A root is a node that has no ancestors, only descendants. A leaf is a node that has no descendants, only ancestors. In a break with convention, I allow that a node may have two parents that are one and the same node if it is connected to that node by two edges. This is a non-standard use of edges, but it will be useful to have two edges connect two nodes such that the child has two parents that are the same node.

6.2.2.2 Arrays of nodes. The rules require drawing nodes in arrays; always, several nodes get drawn at once, and these make up an array. Typically, one array of initial nodes is drawn, and then another array is drawn whose members are attached to members of the first array by directed edges. The first array might represent gametes while the next array represents zygotes, for instance. Edges between two arrays of nodes are always directed toward the nodes that are on the most recently drawn array. Each node on an array is associated with a variable, as is standard in directed acyclic causal graphs. In the graphs I draw, the variable associated with a node is always a relative frequency variable, and it refers either to a haploid individual, a gamete, a zygote, or a mating pair frequency. All nodes on a single array refer to the frequencies of one of these four sorts of individual.

I will refer to the nodes in a graph by the sort of entity to which they refer, so I will talk about haploid nodes, gamete nodes, zygote nodes, and mating pair nodes. Graphs will contain nodes that refer at most to two out of these three sorts of individuals.

Nodes that are members of the same array cannot be attached by edges, though they may be connected by paths. The last array of nodes drawn for a population will represent the relative frequencies of the individuals at the beginning of the subsequent generation, something that I will signal by adding a prime symbol to the relative frequency terms associated with them.

6.2.2.3 Groupings. I will also often find use for the notion of a *grouping*. A grouping is a population, a substructure, or a subgroup, that is, any sort of conglomeration of competitors of the sort considered in selection theory *except* for individuals formed from MICERs, that is, all groupings except gametes, zygotes, and mating pairs.

6.2.2.4 Indexing of nodes. The third step of the algorithm involves the deployment of substitution rules over the relative frequency terms that appear in the equations generated by the graphs at this stage. Accordingly, the relative frequency variables used in the graphs are designed so that *general* substitution rules can be deployed over them, no matter what sort of relative frequency term is being manipulated by substitution. Accordingly, it will be useful to set down rules for indexing the relative frequency variables (always *p* variables) such that the relative frequency terms over which substitution rules are deployed all have indices with identical formats.

The nodes in the graphs will be distinguished by the indices that attach to the relative frequency variables with which they are associated. Generally, this will always be our rule for assigning indices: assign each node a unique index. I specify about how this is done in each case, but the rules for assigning indices to nodes will always be directed at achieving a unique index for each node. We want one node for each type of individual in the population, and we will ascribe indices to relative frequency terms associated with the types of individual represented by the nodes so as to reflect all the possible different types of individual in the system. For the sake of brevity, I will often talk about ascribing indices to the nodes or the individuals they represent, rather than talking about ascribing indices to the relative frequency variables associated with the nodes that pick out the different types of individual in the system.

The approach I use for indexing relative frequency variables builds on one already widespread in population genetics, according to which the relative frequencies of individuals are indexed by *type*, *sex*, and *substructure*. One often sees the variable p_{11} used to refer to organisms that are homozygous for allele A_1. Similarly, one might see the term $p_{22,3}$ used to refer to the relative frequency of individuals who are homozygous for the A_2 allele and live in the third substructure in a hierarchically structured model. One also sees the use of *m* and *f* as subscripts on relative frequency terms to record the sex of the individual bearing a genotype. In the material that follows, each relative frequency term will be accorded a three-part index. The first index will pick out the individual by type, the second by membership in a substructure, and the third by sex; the type, substructure, and sex indices will be separated by commas.

Not every population will include substructures, but by convention, all members of populations that are not hierarchically structured will be treated as living in substructure 1. Haploids and mating pairs do not have sexes, and some diploid organisms lack them, too. Just as we assign everybody to substructure 1 when there are no substructures, we will assign all population members value 0 for their sex index when sex differences are not in play. Following the established practice of the statistical moment approach, we will assign the sex index value 1 for males and value 2 for females in diploid populations. We assign gametes a sex index, too, one reflecting the *sex of origin* of the gametes, so sperm get value 1 and eggs get value 2.

The type index will often be a compound index, one consisting of several elements. Individuals that harbor multiple different alleles, as do multi-locus gametes, all zygotes, and all mating pairs, will all feature compound type indices. Because type indices might otherwise be ambiguous, the following conventions will be used for type indices: Distinct allele types at a locus will appear as distinct numbers, and will be referred to as the *allelic elements* of indices;

angle brackets "<>" will surrounding type indices that differentiate gametes, forming *gametic indices*; square brackets "[]" will surround indices that pick out different types of zygote, forming *zygotic indices*; and braces "{}" will surround indices referring to mating pairs, forming *mating pair indices*. The outermost brackets will not appear in the graphs, however, since the shapes of the nodes that appear in the graphs will signal what sort of individual is being picked out. I use diamond-shaped nodes for gametes to correspond with the angle brackets, square-shaped nodes for zygotes, and hexagons for mating pairs.

In addition, round brackets "()" will surround the alleles at a single genetic locus, forming *genotypic indices*. We will make great use of the notion of a genotypic index in stating the rules that follow. But for now, I will simply point out that the round brackets are necessary to avoid ambiguity. Without the use of the round-brackets, it is not clear whether, say a gamete of type of "<11>" refers to a single-locus gamete with the eleventh allele at one locus or a two-locus gamete with the first allele at both loci. On the convention used here, the first sort of gamete will be ascribed the index <(11)>, while the second will be ascribed the index <(1)(1)>. In models featuring types differentiated in terms of their sex and sex of origin, all indices will be written such that the contribution of a male will be written first and that of the female second (again, following the established practice of the statistical moment approach).

To see how the indexing system works, consider the relative frequency term, $p_{[(12)],1,2}$. This variable represents the relative frequency of a diploid zygote with alleles of interest at a single locus, bearing the A_1 allele inherited from its father and the A_2 allele inherited from its mother, one that is male and living in the second substructure. As another example, the relative frequency of a gamete with allelic variants of interest at two distinct loci, descended from a female, and living in a population without substructures, will be written $p_{<(1)(2)>,2,1}$. The relative

frequency term $p_{\{[(12)(22)(13)][(22)(32)(11)]\},\,2,\,0}$ picks out the frequency of a mating pair, one made up of two diploid zygotes each bearing variant alleles at three loci. The mating pair lives in the second substructure and is sexless. Note that the type indices of mating pairs include allelic elements, genotypic indices, and zygotic indices; they consist in mating pair indices.

6.2.2.5 Genomic imprinting. The equations generated by means of the graphs (though not by means of the algorithm as a whole) will sometimes be too complex for the population at hand. One way the equations might be more complex than necessary is by featuring distinct relative frequency terms for heterozygotes bearing genotypic indices with the same allelic elements, but in different orders, e.g., (12) and (21). In populations with distinct sexes, these two genotypic indices represent different zygotes. The first is formed from a gamete produced by a male and bearing allele A_1 and gamete produced by a female bearing allele A_2, while the second is formed from a gamete produced by a male and bearing allele A_2 and gamete produced by a female bearing allele A_1. However, unless organisms imprint their genomes, the genotypes of heterozygotes that are distinct in this way make causally identical contributions to zygote development.

In equations that are recursions on gamete frequencies, the zygote frequency terms will eventually be replaced when the two sets of equations derived from the graphs are collapsed after the substitution step of the algorithm. So it does not really matter if there are more zygote terms than necessary in the systems of equations that emerge from the lifecycle graphs, since there will not be any unnecessary terms in the equations issued at the end of the algorithm. For equations that are recursions on zygote frequencies, I offer alternative sets of rules for populations with and without genomic imprinting (GI).

6.2.2.6 Gene-by-gene causal influences. For diploid populations, the rules that follow require counting the number of loci that matter to the dynamics of the target alleles. However, no population genetics models feature individuals differentiated by *all* their alleles. That is to be expected, since only sometimes must alleles at multiple loci be taken into account in a classical population genetics model. We have been imagining that the deployment of selection theory by a researcher is triggered by the recognition of the existence of variant competitors, paradigmatically rival alleles. To determine the number of loci that must be considered to accurately model the dynamics of some target competitors, we must ask whether our target alleles causally affected by alleles at other loci. So we can think of the question of whether we need a multi-locus model as a question about what the causal influences of alleles at other loci have to be like such that we are forced to deploy a multi-locus rather than single-locus model. The answer to that question will determine how many loci we must consider.

We need a multi-locus model whenever variant alleles at other loci have either interactive or discriminate causal influences over the genetic variation at the target locus, the one that initially triggered the deployment of the theory. One instance of genetic variation at another locus that is either interactive or discriminate forces a two-locus model, two instances of genetic variations at two other loci that are either interactive or discriminate forces a three-locus model, and so on. I will defend the rule just proposed for cases of discriminate influences and interactive ones emanating from other loci in turn.

That non-neutral alleles at linked loci force the deployment of a multi-locus selection model is a consequence of the fact that their impact on population dynamics evolves as the population evolves. Statistical associations between alleles gradually breakdown among diploids due to recombination among double heterozygotes. Furthermore, such associations may also be

strengthened when individuals that bear alleles in combination have especially elevated relative fitnesses. We cannot just average over the influence of alleles at other loci when these have either a discriminate causal influence on our target alleles, because their impact changes from generation to generation, and the fitness coefficients or functions we use to weight the relative frequencies of our individuals must remain fixed if our models are to fulfill their role of yielding results about long-term evolution. The average fitness individuals with one allele will change as the proportions of individuals bearing that allele and other alleles changes; hence, it cannot be specified by an unchanging coefficient. Fixed relative fitness coefficients (or sometimes functions) are necessary for us to generate analytic results about adaptive evolution, the persistence of polymorphism, and the persistence of altruism.

Non-neutral alleles at other loci that are not statistically associated with our target competitors through linkage need not be taken into account unless their causal influence on individuals' progress interacts with that of our target alleles. That is, even if alleles at some distant locus have a causal influence on the fate of the individuals harboring our target competitors, provided that the alleles at the distant locus have the same influence on each type of our target alleles, their presence need not provoke the deployment of a multi-locus model. That it is possible to overlook indiscriminate *non-interactive* causal influences from alleles at other loci is a result of the mathematical equivalence of models that do and do not do so (Lewontin 1974, 277-78).

The way that interactive or discriminate causal influences of alleles at other loci are handled in classical population genetics is by allowing alleles at those loci to contribute to the differentiation of gametes, zygotes, and mating pairs. When that is done, an allele of one sort can be ascribed one fitness coefficient when paired with one allele at another locus and be assigned a

different relative fitness coefficient when paired with a different allele at another locus. By the inversion principle, the distinct relative fitness coefficients with which individuals are paired reflect the distinct causal influence of the alleles at the other loci. The rates at which these pairings are formed and dissolved are then captured in the model by recombination parameters that specify how statistical associations among the alleles at the different loci change over time. Officially, then, the number of non-pervasive causal influences that are either discriminate or interactive arising from alleles at other loci will determine how we differentiate our individuals in a population genetics model, one whose deployment was triggered by the recognition of variation at a single genetic locus. We differentiate the individuals in a model by the genes they bear at $n + 1$ loci when alleles at n loci have discriminate or interactive causal influences over the alleles at a target locus of interest.

6.2.3 Rules for graph construction

In this section, I begin to lay out the rules for drawing lifecycle graphs, taking advantage of the vocabulary just discussed. While there are rules peculiar to each of the six different sorts of systems distinguished by the decision tree in section 6.1, some rules apply to more than one of the six sorts of systems. I start with rules with especially broad application, and then finish with ones that apply to only one of the six populations distinguished in the previous stage of the algorithm.

6.2.3.1 A general migration rule. I now turn to consider the instructions for drawing lifecycle graphs. I first consider a general migration rule that can be applied to all populations. I then consider haploid graphs, for which the graph-drawing procedure is especially simple. After that I consider models featuring diploids, for which there are once again some general rules.

While the graph-drawing rules for each of the 11 basic sorts of systems considered in what follows are all in some ways different, one fashion in which they are all similar is how migration between substructures is modeled in each. It is possible for both gametes and organisms to travel from one substructure to another, so each graph may feature up to two migration events. Luckily, the rules for drawing a migration event are the same no matter what type of individual one is considering or how many substructures exist in the population, so in the graph-drawing rules that follow, I will simply mention that one must insert a *migration event* into the graph for populations with multiple substructures when it is appropriate to do so, where a migration event will consist in the drawing of an additional array of nodes that is attached to the last ones drawn by directed edges in the fashion I now specify.

A migration event will involve drawing a duplicate array of nodes, nodes that match those of the last array drawn in number and indices. The newly drawn nodes are attached by directed edges to the ones last drawn such that any new node sharing its first and third indices with a pre-migration node is made its child. The edges are then weighted with migration coefficients that are themselves indexed by two numbers, the first number being the value of the second index of the node at the foot of the edge and the second number being the value of the second index of the node at the head of the edge, e.g., m_{12} would weight the edge running from a node representing an individual in the first substructure to a node representing the same type of individual in the second substructure post migration. Note that it is harmless to follow this rule for a non-hierarchically structured population featuring no substructures, since doing so will result in weighting each type by a migration coefficient m_{11} which must take the value 1 since, for migration coefficients, $\sum_i m_{1i} = 1$.

We will see an example of a migration event in the next section. I now turn to consider each of the 11 fundamentally different kinds of systems picked out using the decision-tree earlier on. I begin with haploids.

6.2.3.2 Haploid lifecycle graphs. It is especially simple to graph populations in which competitors do not form MICERs. There is only one sort of relative frequency term in haploid models, unlike the other models we will consider. The initial array of nodes in the graph should be equal to the number of substructures times the number of variant competitors, $Sbs \times A$, where *Sbs* is a variable picking out the number of substructures in our population, and *A* is a variable picking out the number of variant types of haploid individual. The nodes in the graph must be ascribed indexed relative frequency variables that contain information about type and substructure membership. Assign each $1/A$ of the nodes a distinct type index between 1 to *A*. Assign each node a sex index of 0. Assign $1/Sbs$ of the nodes a distinct substructure index between 1 to *Sbs*, such that no two nodes have identical indices. This means that each node has an index that is a permutation of the *A* types of competitor and each of the *Sbs* distinct substructures.

The next thing we need for the haploid lifecycle graph is a depiction of the migration across the substructures, and we show migration in accordance with the general migration rule above. I put that rule into practice here for illustration. To show migration, we must draw another array of nodes, a post-migration array of nodes, equal in number and indexing to the last array drawn. We then draw directed edges between the arrays, with the feet always at the initial nodes and the heads always at the post-migration nodes. In accordance with the migration rule above, draw an edge between a pre-migration node and post-migration one whenever the two share a type index and a sex index. Because all our haploids have the same sex index, 0, this requirement

reduces to the sharing of a type index. Lastly, again in accordance with the general migration rule above, weight these edges with coefficients reflecting migration rates, with parameters m_{xy}, where x is the substructure index of the node at the foot of the edge and y is the substructure index of the node at its head.

Haploids are unusual insofar as we need only one graph to generate a system of equations that can be generalized by substitution to yield a model of the population's dynamics no matter what causal influences operate upon it. Accordingly, we can just treat the post-migration nodes as representing the haploids at the start of the next stage in the lifecycle. Here is an example of a lifecycle graph for a population of haploids of two variant types and three substructures. I use triangles for haploids because they do not look like brackets of any sort.

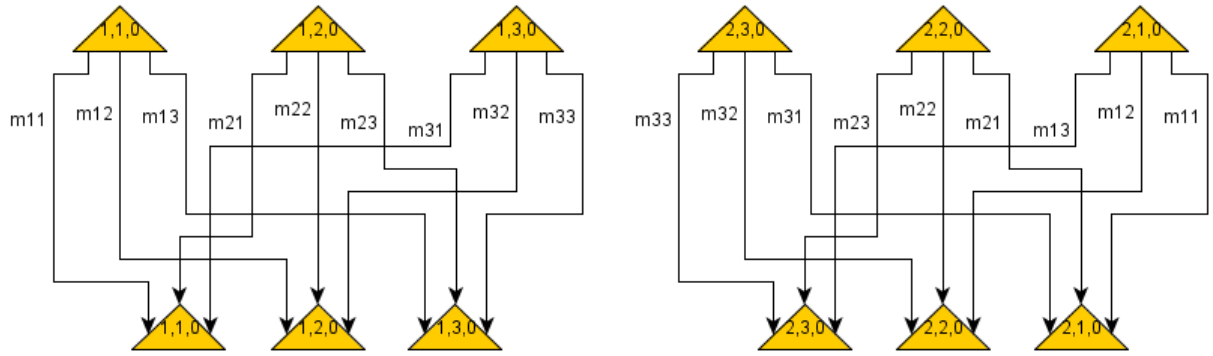


Figure 6.6 Haploid graph with three substructures

The graph yields the following system of equations:

$$p'_{1,1,0} = m_{11}p_{1,1,0} + m_{21}p_{1,2,0} + m_{31}p_{1,1,0}$$
$$p'_{1,2,0} = m_{12}p_{1,1,0} + m_{22}p_{1,2,0} + m_{32}p_{1,1,0}$$
$$p'_{1,3,0} = m_{13}p_{1,1,0} + m_{23}p_{1,2,0} + m_{33}p_{1,1,0}$$
$$p'_{2,1,0} = m_{11}p_{2,1,0} + m_{21}p_{2,2,0} + m_{31}p_{2,1,0}$$
$$p'_{2,2,0} = m_{12}p_{2,2,0} + m_{22}p_{2,2,0} + m_{32}p_{2,1,0}$$
$$p'_{2,3,0} = m_{13}p_{2,3,0} + m_{23}p_{2,2,0} + m_{33}p_{2,1,0}$$

Equations 6.1

These equations will get manipulated by substitution to reflect the causal influences acting on the haploids.

6.2.3.3 Diploid lifecycle graphs. There are five fundamentally different recursive equations for populations in which diploid zygotes are formed: recursions on gametes, recursions on zygotes, sex-dependent versions of these, and recursions suitable for haplo-diploids. Each graphing strategy associated with each of these will be a two-graph affair. The first graph will show either how gametes form zygotes, in case we seek equations that are recursions on gamete frequencies, or how zygotes for mating pairs, in case we seek equations that are recursions on zygote frequencies. The second part of the graph will show how either how zygotes produce next generation gametes, or how mating pairs produce next generation zygotes. The graphs are kept separate so that they yield two sets of equations with distinct relative frequency terms for the distinct types of individuals featured in the graphs.

6.2.3.3.1 General rules for diploid lifecycle graphs. Just as there was a general migration rule that could be deployed whenever migration between substructures occurred, no matter what sort of graph one was composing, there are general rules that can be applied to all diploid graphs, and indeed a handful more that can be applied to subsets of these. I rehearse these here.

6.2.3.3.1.1 Inverted graphs. In all the instructions that follow, the second graph drawn will be the inverse of the first. This means something specific: the second graph will feature the same arrays of nodes indexed in the same fashion, but such that the array of nodes featured as roots in the first graph form the leaves of the second graph, and vice versa. However, nodes are connected by edges differently in each graph, so we will not be able to simply reverse the direction of the edges of the first graph to generate its inversion. Still, we will never have to say how many nodes with what indices are necessary for the second graph; because the second graph is always the inversion of the first, we will only have to say how to connect the nodes in the second graph and we will have specified it in its entirety.

6.2.3.3.1.2 Recombination. The last step in graphical modeling for systems featuring individuals differentiated by the alleles they bear at multiple loci involves weighting the last set of edges in the second graph with recombination rates. Here I state a general rules for weighting these edges so that later I will simply signal that edges must be weighted by recombination parameters. It is to be understood that the rule stated below is to be used to make such weightings.

Recombination may occur as zygotes produce next generation gamete descendants and as mating pairs produce next-generation zygote descendants. The rules are slightly different in each case. While recombination is a fairly simple matter for two-locus, two-allele models, the rules I state for handling recombination are complicated because they are meant to be general and hence should work for systems involving individuals that are differentiated by any number of alleles at any number of loci.

In the graph-drawing rules for multi-locus selection that follow, individuals are distinguished by the alleles they bear at N loci and hence they feature zygotic indices that contain N genotypic indices (recall that these are the round-bracketed entries in the type index that,

among zygotes and mating pairs in which recombination occurs, feature two allelic elements apiece, while featuring one allelic element among gametes). We will ascribe an index 1 to the recombination parameter that expresses the probability of a recombination event occurring between the first and second loci, index 2 to the probability of a recombination event occurring between the second and third loci, index 3 to the probability of a recombination event occurring between the third and fourth loci, and so on, up to N - 1. So, for instance, r_2 is the probability of recombination between the second and third loci.

For models that include more than two loci, we will deploy recombination parameters with compound indices to reflect the probability that recombination occurs between two loci separated by one or more loci. We will have need of these whenever the intermediate loci are ones at which the parent individual is homozygous. So, for instance, if an individual is homozygous at two loci between the first and the fourth loci at which it is a heterozygote, it will not matter whether a single bout of recombination occurs between the first and the second, the second and the third, or the third and the fourth locus. I express the probability of recombination between two distal loci i and $(i + x)$ using the parameter, $r_{i—i+x}$.

The recombination parameter, $r_{i—i+x}$, reflects the probability that there is not a *single*, but rather an *odd number* of recombination events between locus i and locus $(i + x)$. Indeed, it is generally true that r parameters function in this way, such that even r_3 represents the probability that there is an odd number, rather than just a single, recombination event between the third and fourth loci. This means it is not generally true that, say, $r_{2—4}$, is equal to $r_2r_3r_4$. Instead, $r_{2—4}$ will be equal to $r_2r_3r_4 + (1 – r_2)\ (1 – r_3)r_4 + (1 – r_3)(1 – r_4)r_3 + (1 – r_4)(1 – r_3)r_2$. Presumably, facts about recombination rates between any two loci, no matter how distal, will be assessed

statistically such that one need not have a general mathematical function for relating recombination parameters.

I state two sets of rules for weighting edges by recombination parameters. I first consider recombination parameters on edges connecting zygotes to next-generation gametes, and then consider recombination parameters on edges connecting mating pairs to next-generation zygotes.

6.2.3.3.1.2.1 Edge-weighting rule for recombination among zygotes. For zygote parent nodes with non-matching allelic elements in both the i^{th} and the $(i+1)^{th}$ genotypic indices (the parent is a heterozygote at adjacent loci), weight an edge with a parameter to reflect whether recombination has occurred. Weight the edge with recombination parameter r_i provided the (lone) allelic elements of the i^{th} and the $(i+1)^{th}$ genotypic indices of the gamete node's type index do not match both the first, or both the second, allelic elements of the i^{th} and the $(i+1)^{th}$ genotypic places of the node at the foot of the edge; otherwise, weight the edge by a the parameter $(1 - r_i)$.

For zygote parents with non-matching allelic elements at genotypic index i, but not at genotypic indices $(i+1)$ through $(i + x)$ where x is a natural number, such that the parent has non-matching allelic indices at the i^{th} and the $(i + x + 1)^{th}$ genotypic indices (the zygote parent is homozygous at one or more loci between two loci at which it is heterozygous), weight the edge with a parameter reflecting whether recombination has occurred between the distal loci. Weight the edge by parameter $r_{i—i+x}$ provided the (lone) allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the gamete node at the head of the edge do not match both the first, or both the second, allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the node at the foot of the edge; otherwise, weight the edge by a the parameter $(1 - r_{i—i+x})$.

6.2.3.3.1.2.2 Edge-weighting rule for recombination among mating pairs. For mating pair parent nodes with non-matching allelic elements in both the i^{th} and the $(i+1)^{th}$ genotypic indices

of their first zygotic index, weight an edge with a parameter to reflect whether recombination has occurred. Weight the edge with recombination parameter r_i provided the first allelic elements of the i^{th} and the $(i + 1)^{th}$ genotypic indices of the descendant zygote node's type index do not match both the first, or both the second, allelic elements of the i^{th} and the $(i + 1)^{th}$ genotypic indices of the first zygotic index of the node at the foot of the edge; otherwise, weight the edge by the parameter $(1 - r_i)$. Similarly, for mating pair parent nodes with non-matching allelic elements in both the i^{th} and the $(i + 1)^{th}$ genotypic indices of their second zygotic index, weight an edge with a parameter to reflect whether recombination has occurred. Weight the edge with recombination parameter r_i provided the second allelic elements of the i^{th} and the $(i + 1)^{th}$ genotypic indices of the descendant zygote node's type index do not match both the first, or both the second, allelic elements of the i^{th} and the $(i + 1)^{th}$ genotypic indices of the second zygotic index of the node at the foot of the edge; otherwise, weight the edge by a the parameter $(1 - r_i)$.

For mating pair parent nodes with non-matching allelic elements in their first zygotic index at genotypic index i, but not at genotypic indices $(i + 1)$ through $(i + x)$ where $x \geq 1$, such that the parent has non-matching allelic indices at the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of its first zygotic index, weight the edge with a parameter reflecting whether recombination has occurred between the distal loci. Weight the edge by parameter $r_{i—i+x}$ provided the first allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the zygote node at the head of the edge do not match both the first, or both the second, allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the first zygotic index of the node at the foot of the edge; otherwise, weight the edge by the parameter $(1 - r_{i—i+x})$. Similarly, for mating pair parent nodes with non-matching allelic elements in their second zygotic index at genotypic index i, but not at genotypic index $(i + 1)$ through $(i + x)$ where $x \geq 1$, such that the parent has non-matching allelic indices at the i^{th}

and the $(i + x + 1)^{th}$ genotypic indices of their second zygotic index, weight the edge with a parameter reflecting whether recombination has occurred between the distal loci. Weight the edge by parameter $r_{i—i+x}$ provided the second allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the zygote node at the head of the edge do not match both the first, or both the second, allelic elements of the i^{th} and the $(i + x + 1)^{th}$ genotypic indices of the first zygotic index of the node at the foot of the edge; otherwise, weight the edge by a the parameter $(1 - r_{i—i+x})$.

6.2.3.3.1.3 Meiosis and meiotic drive. We must also weight the last set of edges in the second graph of diploid organisms using parameters reflecting meiosis. We must equally weight them with parameters expressing meiotic drive, when it occurs. The strategy pursued here is to weight every edge emerging from heterozygotic parents with a parameter reflecting unbiased meiosis, and then to further weight edges emerging from heterozygotes bearing driving alleles to reflect meiotic drive. The rules are simpler when stated in this fashion, but the strategy requires that the parameters that quantify the extent to which driving alleles and their partners at the same locus are favored or disfavored in the process of fertilization will vary from 0 to 2, rather than 0 to 1 as is standard. The drive parameters used here take on a greater range of quantities because they are used to weight edges already weighted by a coefficient of ½ that represents unbiased meiosis; standard meiotic drive functions replace, rather than weight, parameters expressing unbiased meiosis. The approach to modeling meiotic drive and the approach pursued here are mathematically equivalent.

Edges that are candidates for being weighted to reflect meiosis all emerge from the same array. Assign an edge a parameter reflecting meiosis provided it emerges from a node in the array bearing any genotypic index with non-matching allelic elements. Weight the edge by a value of $1/2c$, where c is the number of genotypic indices with non-matching allelic elements in the type index of the node at the head of the edge. Thus, zygotes that are heterozygotic at one locus will produce gametes bearing one of their two variant alleles, double heterozygotes will produce four types of gametes, and so on.

To quantify meiotic drive, weight the edges emerging from a zygote bearing a driving allele with functions, k's. The k parameters should feature two indices in sequence separated by a comma, the second reflecting which of the N loci the allele is driving. The first index should consist in two numbers, the first number is the index of the allele passed on to the descendant, and the second is the other allelic element at the same locus in the parent. So, for instance, $k_{12,\,2}$, would represent a drive parameter governing how effectively the A_1 allele competes against the A_2 allele at the locus 2, whenever the A_1 allele is passed on.

Meiotic drive parameters are symmetrical, so that, for any x and y, $k_{xy} + k_{yx} = 2$. The more the meiotic drive parameters deviate from 1, the more the driving allele biases zygote formation. Driving alleles will usually bias zygote formation to the same extent no matter what the allele with which they are paired, so that k_{xy} will equal k_{xz} for all y and z, but the indexing system we have used does not force us to assume this.

6.2.3.3.1.4 Discriminate union of individuals. Another set of general rules for diploid graphs can be stated upfront, one for weighting edges by parameters that express the extent to which the union of individuals into higher order individuals is discriminate. The general term, "discriminate union of individuals," covers assortative mating, selfing, gametic self-

incompatibility, and any other form of discriminate pairing among individuals in the system. Gametes may form zygotes discriminately, such that they are more likely to pair with other gametes of one type rather than another. Similarly, zygotes may form mating pairs with other zygotes discriminately. Both these sorts of facts are captured by weighting the edges of lifecycle graphs that represent the formation of such unions. There is a general rule for making such weightings, and once again I will state it here and then later simply give the instructions to weight the edges by parameters reflecting the discriminate union of individuals when appropriate.

Edges must be weighted by parameters expressing the extent to which individuals pair discriminately whenever individuals engage in MICERs discriminately. Edges emerging from nodes picking out individuals that pair discriminately must each be weighted by parameters, a_{xy}, where x is the type index of individual at the foot of the edge, and y is the type index of the other parent of the node at the head of the edge (the other member of the pair), and a expresses the extent to which one type of pairing is preferred as compared to the others. Discriminate union parameters are subject to the constraint that $\sum_{y}(a_{xy} \times p_{y}) = 1$, where p_y is the relative frequency of the individual with which the pairings are formed.

In a footnote, I earlier claimed that sexual selection was not an aspect of selection in the same sense that gametic, zygotic, and fecundity selection were aspects of selection. Here is my reason: capturing "sexual selection" is a matter of deploying parameters that represent the extent to which zygotes pair with other zygotes in a discriminate fashion. We will later see that the other three of the traditional four aspects of selection represent the impact of ecological causal influences over the progress of each type of individual to the next stage in the lifecycle. These are very different sorts of things handled very differently.

6.2.3.3.2 Sex-dependent selection graphs. Having laid down a set of rules to which we can appeal when composing graphs for populations featuring diploids, we now come to consider a subset of these, the four sorts of graphs that feature individuals differentiated by sex and sex of origin. There is a general rule for drawing graphs under sex-dependent selection and another for sex-independent selection. For sex-dependent graphs, it is possible to state a fairly simple general rule for connecting arrays of nodes representing sex-of-origin differentiated gametes to an array of nodes representing sex-differentiated zygotes. A very similar rule works for connecting nodes that represent sex-differentiated zygotes to ones that represent mating pairs that are not differentiated by sex (since mating pairs have no sexes). So provided we have indexed the nodes in the initial array of nodes in the first graph, we can connect them to the nodes in the second array we draw following a simple rule. What's more, we can index the nodes in the second array following a simple rule too, by compounding the indices of their parents. I call the rule for performing these actions the *sex-dependent connect-and-compound rule*.

6.2.3.3.2.1 Sex-dependent connect-and-compound rule. Join each initial node with a sex index of 1 to every other node in the array of initial nodes which it shares its substructure index, but not its sex index, by a two-edge path through a secondary node that is so far unattached to any other node. Both edges will point from the initial nodes to the secondary nodes, so following a path will involve "following" an edge from head to foot. If the secondary nodes that are now leaves in the graph are mating pairs, then the edge-drawing procedure is finished. If they are zygotes, assign each secondary node so far connected to the initial nodes a sex index of one (these are the male zygotes). Then repeat the path-drawing procedure just deployed for the remaining secondary nodes, attaching each initial node with sex index one to another initial node with sex index of two that shares its substructure index by a two-edge path through a unique so far

unattached secondary node. This should exhaust the secondary nodes such that each has exactly two parents. Assign the remaining secondary nodes without sex indices a sex index of two (these are the female zygotes). Assign each secondary node the same substructure index of its parents.

Now we assign each secondary node a type index that is the compound of the type indices of its parents. Zygote type indices are compounded to form mating pair indices quite simply. One simply places the type indices of the parent nodes side by side between a set of braces, with the type index of the node with sex index 1 first (the male) and the type index of the parent node with sex index 2 second (the female). The resulting bracketed term is the mating pair index for the node. So, for instance, if one parent has type index, [(12)(22)], and the other has type index, [(22)(33)], the mating pair they form will have type index {[(12)(22)][(22)(33)]}. Mating pairs always have sex index zero, and children always inherit the substructure index of their parents.

Compounding gamete indices to form zygote indices is a little more complicated. To generate the type index for a zygote node from the gamete indices of its two parents, pair the allelic elements of each of the *k* genotypic indices of the gamete nodes with the allelic element of the gamete node with sex index coming one first, to generate an identical number of two-place genotypic indices. String the new genotypic indices together, maintaining the order such that the first allelic element of the *k*th genotypic index just generated matches the sole allelic element of the *k*th genotypic index of the parent gamete node with sex index one, and the second allelic element of the *k*th genotypic index just generated matches the sole allelic element of the *k*th genotypic index of the parent gamete node with sex index two. Place the strings of genotypic indices within square brackets to yield the type index for the zygote. Once again, the zygote child must share the substructure index of its parents. So, for instance, two gamete nodes whose

indices are these: <(1)(2)>, 1, 2 and <(2)(3)>, 2, 2 will produce two sorts of children: [(12)(23)], 1, 2; and [(12)(23)], 2, 2.

6.2.3.3.2.2 Graphing instructions for sex-dependent recursions on gamete frequencies.

Draw $2 \times Sbs \times \prod_{k=1}^{N} A_k$ nodes to represent the gametes picked out by the alleles they bear at multiple loci, where A_k is the number of alleles at locus k of N total loci ($k = 1 \ldots N$), *Sbs* is the number of substructures, and the coefficient of two represents the two sexes. We now seek to assign each node a unique index. To do so, first assign each $1/Sbs$ of the gamete nodes a substructure index from 1 to *Sbs*. Assign half of the nodes with each substructure index sex index 1 and the other half sex index 2. Next, assign nodes indices such that each gamete node features N genotypic places with room for a single allelic element. There are A_k possibilities for the allelic elements at each kth genotypic place. Assign each $1/A_1$ of the nodes sharing the same sex and substructure indices each value from 1 to A_1 to its first genotypic place. Assign $1/A_2$ of the nodes each value from 1 to A_2 to its second genotypic place, and so forth. Generally, assign $1/A_x$ of the nodes sharing sex and substructure indices each value from 1 to A_x to its x^{th} genotypic place for $x = 1$ to N, such that no two nodes sharing a substructure index and a sex also share a type index, or, more succinctly, such that no two nodes share an index. Add a migration event here for populations with multiple substructures and mobile gametes.

The number of zygote nodes in the lifecycle graph is fixed by this formula: $2 \times Sbs \times \left(\prod_{k=1}^{N} A_k \right)^2$, twice the number of ways the distinct male and female gametes from each substructure can be permuted to form zygotes. Attach the zygote nodes to the gamete nodes using the connect-and-compound rule for sex-dependent selection. For cases in which gametes form zygotes in a discriminate fashion, the edges just drawn must be weighted by parameters

expressing the differential rates at which zygotes are formed according to the discriminate union of individuals rule. A migration event is now appropriate for populations with mobile zygotes and multiple substructures.

Once again, we need only consider how to connect the nodes in the second graph because it is the inversion of the first. A gamete node should be connected to zygote node provided that the following conditions hold:

- the two nodes share a substructure index and a sex index, and
- the allelic elements in each of the *k* genotypic indices of the type index of the secondary nodes matches at least one of the allelic elements in each of the corresponding *k* genotypic indices of the initial nodes.

The edges must now be weighted by coefficients representing meiosis, meiotic drive, and recombination rates, using the appropriate rules.

Here is an example of a simple single-locus lifecycle graph featuring two alleles and no substructures for sex-dependent selection. Here is the first graph:

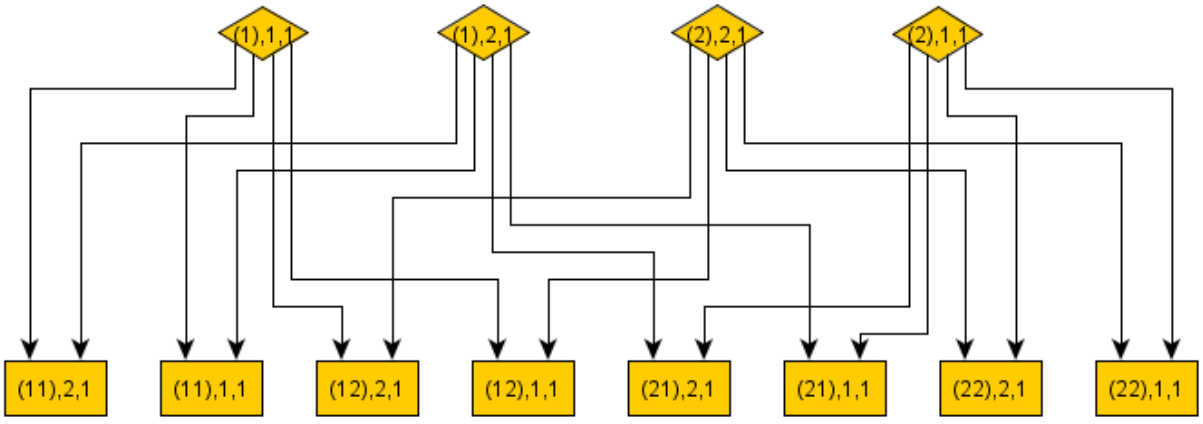


Figure 6.7 Single-locus sex-dependent selection graph one

It yields the following set of equations:

$$p_{[(11)],2,1} = p_{<(1)>,1,1}p_{<(1)>,2,1}$$

$$p_{[(11)],1,1} = p_{<(1)>,1,1}p_{<(1)>,2,1}$$

$$p_{[(12)],2,1} = p_{<(1)>,1,1}p_{<(2)>,2,1}$$

$$p_{[(12)],1,1} = p_{<(1)>,1,1}p_{<(2)>,2,1}$$

$$p_{[(21)],2,1} = p_{<(2)>,1,1}p_{<(1)>,2,1}$$

$$p_{[(21)],1,1} = p_{<(2)>,1,1}p_{<(1)>,2,1}$$

$$p_{[(22)],2,1} = p_{<(2)>,1,1}p_{<(2)>,2,1}$$

$$p_{[(22)],1,1} = p_{<(2)>,1,1}p_{<(2)>,2,1}$$

Equations 6.2

Because the graph shows the formation of zygotes from gametes, descendant variables are represented as the mathematical *products* of their ancestors. Note also how I use the shape of the node to signal what sorts of individual is being picked out; the diamond-shaped nodes are gametes, and the square ones zygotes. Here is the second graph:

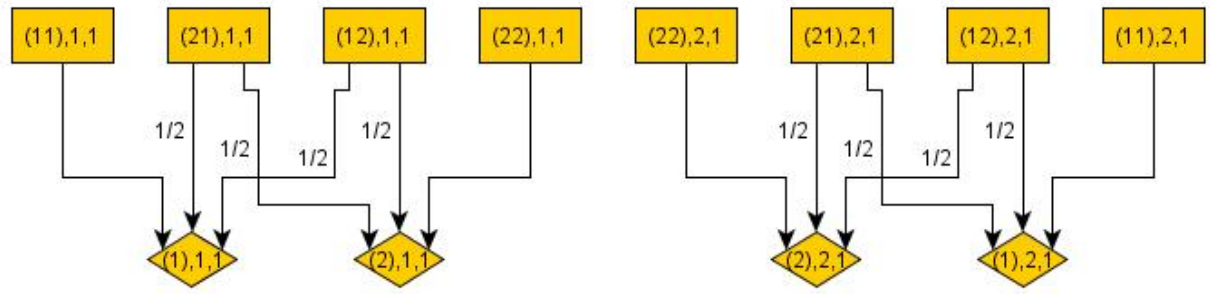


Figure 6.8 Single-locus sex-dependent selection graph two

Note that the edges in the second graph are not the inverses of the edges in the first graph, though the arrays of nodes are inverted duplicates of the ones in the first graph. Here are the equations we can infer from the second graph:

$$p'_{<(1)>,1,1} = p_{[(11)],1,1} + \frac{1}{2} p_{[(21)],1,1} + \frac{1}{2} p_{[(12)],1,1}$$
$$p'_{<(2)>,1,1} = \frac{1}{2} p_{[(21)],1,1} + \frac{1}{2} p_{[(12)],1,1} + p_{[(22)],1,1}$$
$$p'_{<(1)>,1,1} = p_{[(22)],2,1} + \frac{1}{2} p_{[(21)],2,1} + \frac{1}{2} p_{[(12)],2,1}$$
$$p'_{<(1)>,2,1} = \frac{1}{2} p_{[(21)],2,1} + \frac{1}{2} p_{[(12)],2,1} + p_{[(11)],2,1}$$

Equations 6.3

Because the second graph shows the breakdown of zygotes into gametes, the descendant variables are the sums of their ancestors.

I propose to jump ahead a little here and show how we can get the textbook model of sex-dependent selection if we 1) weight the zygote frequencies by selection coefficients, and 2) collapse the systems of equations. Officially, we have not yet learned the rules for doing this, but it might be interesting to see how to finish things off in this particular case. This rest section can be skipped by the somewhat incurious (presumably the truly incurious are not even reading this).

Weighting the zygote relative frequency parameters in the second set of equations by relative fitness and average fitness parameters, where $\overline{w}$ is the average zygote fitness, we get this result:

$$p'_{<(1)>,1,1} = \frac{w_{[(11)],1,1}}{\overline{w}} p_{[(11)],1,1} + \frac{1}{2} \frac{w_{[(21)],1,1}}{\overline{w}} p_{[(21)],1,1} + \frac{1}{2} \frac{w_{[(12)],1,1}}{\overline{w}} p_{[(12)],1,1}$$
$$p'_{<(2)>,1,1} = \frac{1}{2} \frac{w_{[(21)],1,1}}{\overline{w}} p_{[(21)],1,1} + \frac{1}{2} \frac{w_{[(12)],1,1}}{\overline{w}} p_{[(12)],1,1} + \frac{w_{[(22)],1,1}}{\overline{w}} p_{[(22)],1,1}$$
$$p'_{<(2)>,2,1} = \frac{w_{[(22)],2,1}}{\overline{w}} p_{[(22)],2,1} + \frac{1}{2} \frac{w_{[(21)],2,1}}{\overline{w}} p_{[(21)],2,1} + \frac{1}{2} \frac{w_{[(12)],2,1}}{\overline{w}} p_{[(12)],2,1}$$
$$p'_{<(1)>,2,1} = \frac{1}{2} \frac{w_{[(21)],2,1}}{\overline{w}} p_{[(21)],2,1} + \frac{1}{2} \frac{w_{[(12)],2,1}}{\overline{w}} p_{[(12)],2,1} + \frac{w_{[(11)],2,1}}{\overline{w}} p_{[(11)],2,1}$$

Equations 6.4

Since the only kind of selection we will consider is differential viability of zygotes, we can now collapse the equations by using the first set of equations to fix functions for the corresponding right-hand side variables of the second set. This yields a set of equations with variables that refer only to gametes and hence the equations are recursions on gamete frequencies:

$$p'_{<(1)>,1,1} = \frac{w_{[(11)],1,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(1)>,2,1} + \frac{1}{2}\frac{w_{[(21)],1,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(1)>,2,1} + \frac{1}{2}\frac{w_{[(12)],1,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(2)>,2,1}$$

$$p'_{<(2)>,1,1} = \frac{1}{2}\frac{w_{[(21)],1,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(1)>,2,1} + \frac{1}{2}\frac{w_{[(12)],1,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(2)>,2,1} + \frac{w_{[(22)],1,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(2)>,2,1}$$

$$p'_{<(2)>,2,1} = \frac{w_{[(22)],2,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(2)>,2,1} + \frac{1}{2}\frac{w_{[(21)],2,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(1)>,2,1} + \frac{1}{2}\frac{w_{[(12)],2,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(2)>,2,1}$$

$$p'_{<(1)>,2,1} = \frac{1}{2}\frac{w_{[(21)],2,1}}{\overline{w}} p_{<(2)>,1,1} p_{<(1)>,2,1} + \frac{1}{2}\frac{w_{[(12)],2,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(2)>,2,1} + \frac{w_{[(11)],2,1}}{\overline{w}} p_{<(1)>,1,1} p_{<(1)>,2,1}$$

Equations 6.5

If we set the relative fitness values of the heterozygotes equal, as is appropriate for populations in which organisms do not engage in sex-dependent genomic imprinting at the locus in equation, then the first and fourth equations are the standard equations for the change in frequency of the A_1 allele under biallelic sex-dependent selection (Ewens 2004, 46). The second and third equations are redundant, since $p'_{<(2)>,1,1} = 1 - p'_{<(1)>,1,1}$ and $p'_{<(2)>,2,1} = 1 - p'_{<(1)>,2,1}$, that is, the relative frequency of the A_2 allele in males and females is just one minus the relative frequency of the A_1 allele in each sex-context because there are only two alleles in the population.

6.2.3.3.2.3 Graphing instructions for sex-dependent recursions on zygote frequencies. The number of zygote nodes in the first graph is equal to $2 \times Sbs \times \left(\prod_{k=1}^{N} A_k \right)^2$. For cases without genomic imprinting, the formula is this: $2 \times Sbs \times \prod_{k=1}^{N} \sum_{i=1}^{A_k} i$.

Assign 1/*Sbs* of the nodes a unique substructure index ranging from 1 to *Sbs*. Assign half of the nodes in each substructure a sex index of 1 and the other half a sex index of 2. The type index of each node will consist in a square-bracketed zygotic index containing *N* genotypic places featuring for two allelic places apiece. To generate the type indices, permute (if no GI, combine) the alleles at each of the A_k loci to generate variant genotypic indices at each locus. Permute these variant genotypic indices with each other to generate variant zygotic indices. Assign the various zygotic indices to the zygote nodes such that no two nodes have the same index. A migration event is appropriate here for populations with multiple substructures and mobile zygotes.

Draw $Sbs \times \left(\prod_{k=1}^{N} A_k \right)^4$ nodes to represent the mating pairs. For cases without genomic imprinting, the formula is this: $Sbs \times \left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right)^2$. To attach these nodes to the zygote nodes, follow the connect-and-compound rule. The edges just generated must be weighted by assortative mating parameters for whichever zygotes form mating pairs discriminately by using the discriminate union of individuals rule.

In the second graph, a zygote node must be made the descendant of a mating pair node if the zygote node shares its substructure index with the mating pair node and the following two conditions hold:

- one of the two allelic elements of each of the *k* genotypic indices of the first zygotic index of the mating pair node matches the first (if no GI: either of the) allelic element of each of the corresponding *k* genotypic indices of the type index of the secondary node, and

- one of the two allelic elements of each of the k genotypic indices of the second zygotic index of the initial node matches the second (if no GI: either of the) allelic element of each of the corresponding k^{th} genotypic indices of the type index of the secondary node.

These edges must be weighted to reflect recombination, meiosis, and parameters reflecting meiotic drive for any driving alleles.

Here is an example of the two lifecycle graphs appropriate for a single-locus two-allele recursion on zygote frequencies in which females, but not males, mate assortatively, mating pairs produce offspring of different sexes in equal proportions, and there is neither genomic imprinting nor meiotic drive:

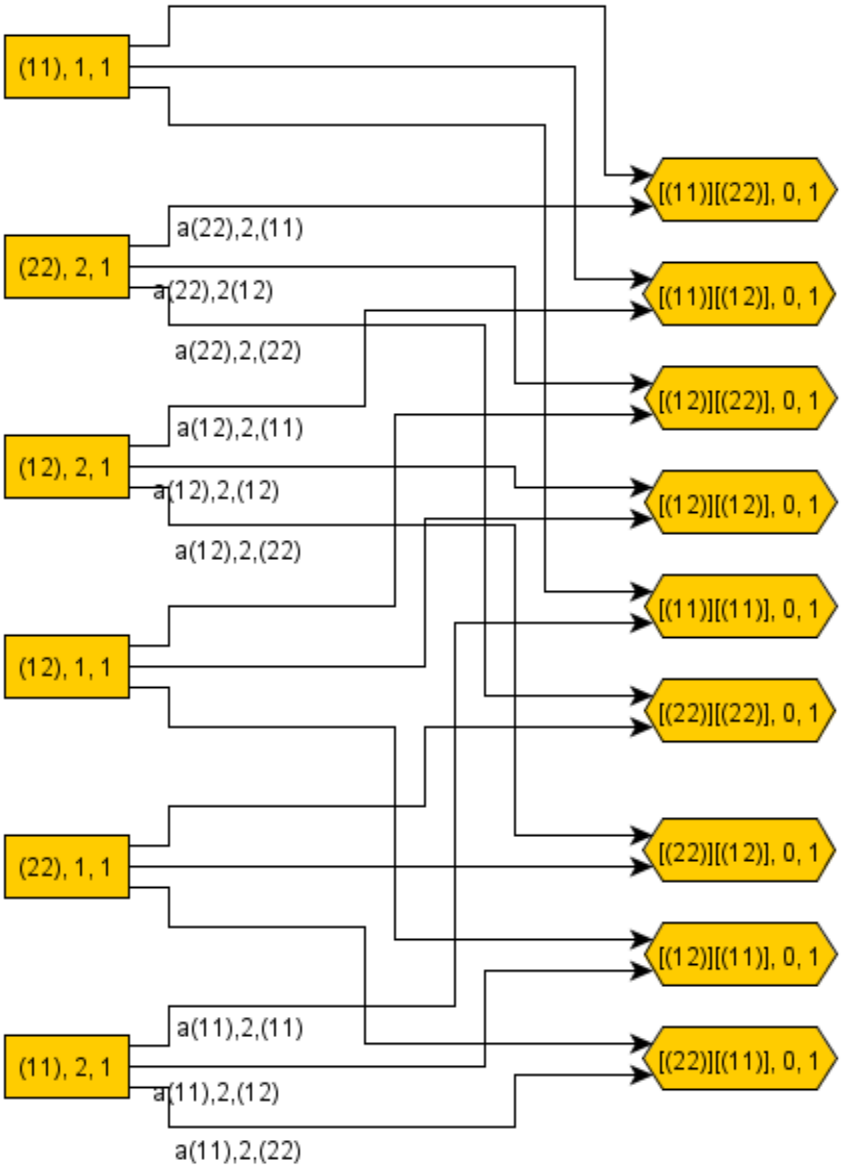


Figure 6.9 Graph one of population with female assortative mating

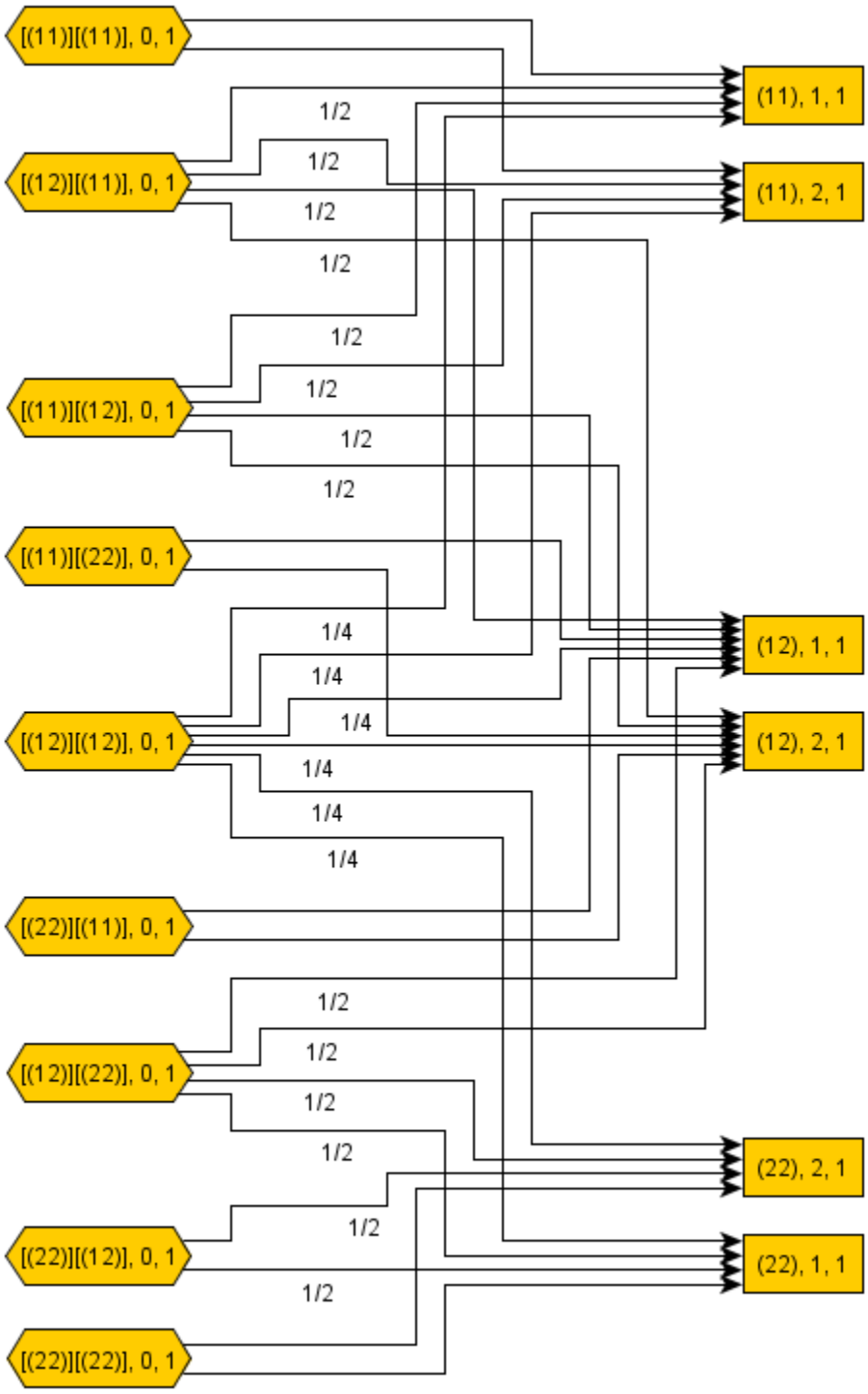


Figure 6.10 Graph two of population with female assortative mating

The graphs yield these two systems of equations:

$$p_{\{[(11)][(22)]\},0,1} = p_{[(11)],1,1}a_{(22),2,(11)}p_{[(22)],2,1}$$
$$p_{\{[(11)][(12)]\},0,1} = p_{[(11)],1,1}a_{(12),2,(11)}p_{[(12)],2,1}$$
$$p_{\{[(12)][(22)]\},0,1} = p_{[(12)],1,1}a_{(22),2,(12)}p_{[(22)],2,1}$$
$$p_{\{[(12)][(12)]\},0,1} = p_{[(12)],1,1}a_{(12),2,(12)}p_{[(12)],2,1}$$
$$p_{\{[(11)][(11)]\},0,1} = p_{[(11)],1,1}a_{(11),2,(11)}p_{[(11)],2,1}$$
$$p_{\{[(22)][(22)]\},0,1} = p_{[(22)],1,1}a_{(22),2,(22)}p_{[(22)],2,1}$$
$$p_{\{[(12)][(11)]\},0,1} = p_{[(12)],1,1}a_{(11),2,(12)}p_{[(11)],2,1}$$
$$p_{\{[(22)][(12)]\},0,1} = p_{[(22)],1,1}a_{(12),2,(22)}p_{[(12)],2,1}$$
$$p_{\{[(22)][(11)]\},0,1} = p_{[(22)],1,1}a_{(11),2,(22)}p_{[(11)],2,1}$$

Equations 6.6

and

$$p'_{[(11)],1,1} = p_{\{[(11)][(11)]\},0,1} + \frac{1}{2}p_{\{[(12)][(11)]\},0,1} + \frac{1}{2}p_{\{[(11)][(12)]\},0,1} + \frac{1}{4}p_{\{[(12)][(12)]\},0,1}$$
$$p'_{[(11)],2,1} = p_{\{[(11)][(11)]\},0,1} + \frac{1}{2}p_{\{[(12)][(11)]\},0,1} + \frac{1}{2}p_{\{[(11)][(12)]\},0,1} + \frac{1}{4}p_{\{[(12)][(12)]\},0,1}$$
$$p'_{[(12)],1,1} = \frac{1}{2}p_{\{[(12)][(11)]\},0,1} + \frac{1}{2}p_{\{[(11)][(12)]\},0,1} + p_{\{[(11)][(22)]\},0,1} + \frac{1}{4}p_{\{[(12)][(12)]\},0,1}$$
$$+ p_{\{[(22)][(11)]\},0,1}\frac{1}{2}p_{\{[(12)][(22)]\},0,1}$$
$$p'_{[(12)],2,1} = \frac{1}{2}p_{\{[(12)][(11)]\},0,1} + \frac{1}{2}p_{\{[(11)][(12)]\},0,1} + p_{\{[(11)][(22)]\},0,1} + \frac{1}{4}p_{\{[(12)][(12)]\},0,1}$$
$$+ p_{\{[(22)][(11)]\},0,1}\frac{1}{2}p_{\{[(12)][(22)]\},0,1}$$
$$p'_{[(22)],2,1} = \frac{1}{4}p_{\{[(12)][(12)]\},0,1} + \frac{1}{2}p_{\{[(12)][(22)]\},0,1} + \frac{1}{2}p_{\{[(22)][(12)]\},0,1} + p_{\{[(22)][(22)]\},0,1}$$
$$p'_{[(22)],1,1} = \frac{1}{4}p_{\{[(12)][(12)]\},0,1} + \frac{1}{2}p_{\{[(12)][(22)]\},0,1} + \frac{1}{2}p_{\{[(22)][(12)]\},0,1} + p_{\{[(22)][(22)]\},0,1}$$

Equations 6.7

As an aside, I note that if we let the assortative mating parameters take values according to the following functions:

$$a_{11,2,11} = (1-s)/C_{11}$$
$$a_{11,2,12} = (1-hs)/C_{11}$$
$$a_{11,2,22} = 1/C_{11}$$
$$a_{12,2,11} = (1-hs)/C_{12}$$
$$a_{12,2,12} = (1-s)/C_{12}$$
$$a_{12,2,22} = (1-hs)/C_{12}$$
$$a_{22,2,11} = 1/C_{22}$$
$$a_{22,2,12} = (1-hs)/C_{22}$$
$$a_{22,2,11} = (1-s)/C_{22}$$

Equations 6.8

where C_{xy} is a correction factor to ensure that the assortative mating parameters sum to one when weighted by mate frequencies, we get a model of the dynamics of the human MHC complex genes if we make no further modifications and collapse the equations immediately (Hedrick 2005, 196). Setting the assortative mating parameters in the above way reflects a disposition on the part of zygotes to mate with others who share especially few alleles with them at the locus in question. Human females have been found in some studies to be especially attracted to individuals who share few MHC alleles with them based on their odor, as diagnosed through t-shirt tests; just the opposite occurs with women taking birth control pills (Wedekind et al. 1995).

6.2.3.3.3 Sex-independent selection graphs. We now consider the ones in which sex differences are not officially registered. For all these models, the sex index of every node is set to zero. There is a connect-and-compound rule for models without sex differentiation, too.

6.2.3.3.3.1 Sex-independent connect-and-compound rule. The rule for connecting the nodes in the first array to the nodes in the second array is simple: Join each initial node to every other

node in the array of initial nodes which it shares its substructure index, *including itself*, by a two-edge path through a secondary node that is so far unattached to any other node.

If the children are nodes representing mating pairs, assign them type indices in braces consisting of the two type indices of their parents strung together in any order. If the children are nodes representing zygotes, assign the zygote node a square-bracketed type index featuring as many genotypic indices, N, as are used to differentiate the gametes. Assign two allelic elements to each k of the N genotypic indices of the type index of the zygote, the first matching the allelic element of the k^{th} genotypic index of one parent, and the second matching the allelic element of the k^{th} genotypic index of the other parent, such that the value of the first allelic element is also less than or equal to the value of the second allelic element. (That last constraint is arbitrary, but some constraint must be put in place to avoid confusion, and traditionally heterozygotes are picked out using indices such as p_{12} rather than p_{21}.) Once again, the children must share the substructure index of their parents and have sex index zero.

6.2.3.3.3.2 Graphing instructions for sex-independent recursions on gamete frequencies.

The number of gamete nodes can be expressed by the following formula: $Sbs \times \prod_{k=1}^{N} A_k$, where A_k is the number of alleles at locus k, and Sbs is the number of substructures. Assign each $1/Sbs$ of the initial nodes a substructure index from 1 to Sbs. The type indices will have to feature N genotypic places with A_k variant alleles at each locus k. Assign $1/A_x$ of the nodes each value from 1 to A_x to its xth genotypic place for $x = 1$ to N. Do this such that no two nodes sharing a substructure index also share a type index, or, more succinctly, such that no two nodes share an index.

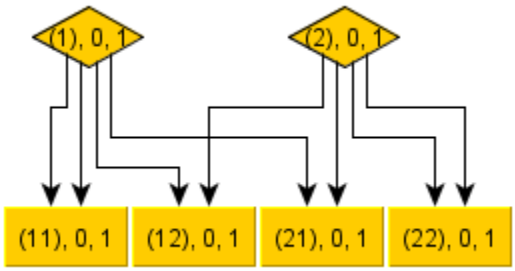
The number of zygote nodes is equal to $Sbs \times \sum_{1}^{i} i; i = \left(\prod_{k=1}^{N} A_k \right)$. Join the gamete nodes to the zygote nodes using the connect-and-compound rule. For cases in which gametes form zygotes in a discriminate fashion, the edges just drawn must be weighted by parameters expressing the differential rates at which pair zygotes are formed using the discriminate union of individuals rule. A migration event is now appropriate for populations with mobile zygotes and multiple substructures.

Connect a zygote node to a gamete node provided that

- the two nodes share a substructure index, and
- the allelic elements in each of the 1 to *N* genotypic indices of the type index of the secondary node match at least one of the allelic elements in each of the corresponding 1 to *N* genotypic indices of the initial nodes.

These edges must now be weighted to reflect recombination, meiosis, and, if necessary, meiotic drive, all in accordance with the general rules for doing so.

Here a simple example of a two-allele model featuring meiotic drive:



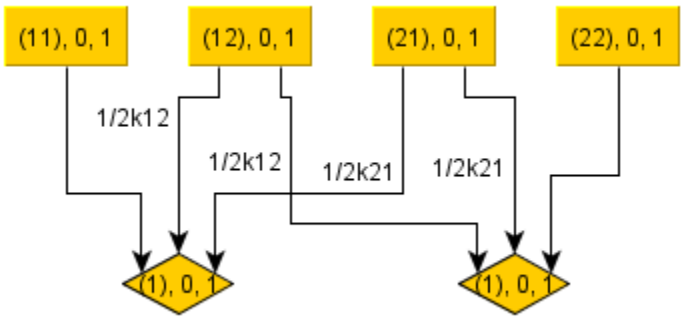


Figure 6.11 Graph of population with meiotic drive

The equations we get from the first lifecycles graph are these:

$$p_{[(11)],0,1} = p_{<(1)>,0,1} \times p_{<(1)>,0,1}$$
$$p_{[(12)],0,1} = p_{<(1)>,0,1} \times p_{<(2)>,0,1}$$
$$p_{[(21)],0,1} = p_{<(1)>,0,1} \times p_{<(2)>,0,1}$$
$$p_{[(22)],0,1} = p_{<(2)>,0,1} \times p_{<(2)>,0,1}$$

Equations 6.9

The equations we get from the second lifecycle graph are these:

$$p_{<(1)>,0,1} = p_{[(11)],0,1} + \frac{1}{2} p_{[(12)],0,1} k_{12} + \frac{1}{2} p_{[(21)],0,1} k_{21}$$
$$p_{<(2)>,0,1} = p_{[(22)],0,1} + \frac{1}{2} p_{[(12)],0,1} k_{12} + \frac{1}{2} p_{[(21)],0,1} k_{21}$$

6.2.3.3.2.3 Graphing instructions for sex-independent recursions on zygote frequencies. The number of zygote nodes in the lifecycle graph is equal to $Sbs \times \left(\prod_{k=1}^{N} A_k \right)^2$, where *Sbs* is substructure number, and A_k is the number of alleles at each of the *k* loci of interest. For populations without GI, the formulation is $Sbs \times \left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right)$. As usual, these nodes must be indexed such that each has a unique index. First assign 1/*Sbs* of the nodes a unique substructure index ranging from 1 to *Sbs*. Then assign half of the nodes in each substructure a sex index of 1

and the other half a sex index of 2. Lastly, each node must be assigned a type index. Each type index will consist of a square-bracketed zygotic index containing N genotypic indices featuring two allelic elements apiece. The idea is once again to assign to each node one every possible index. To generate these, permute (if no GI: combine) the alleles at each of the A_k loci to generate variant genotypic indices. Permute these with each other to generate variant zygotic indices. Then assign the zygotic indices to the initial nodes such that no two nodes have the same index.

The secondary array of nodes is used to represent mating pairs. Draw $Sbs \times \left(\prod_{k=1}^{N} A_k \right)^4$ nodes to represent the mating pairs. For populations without genomic imprinting, the formula is $Sbs \times \left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right)^2$. To attach the mating pair nodes to the zygote nodes and index the former, use the connect-and-compound rule.

A zygote node is made the child of a mating pair node if the zygote node shares its second index with the mating pair node and

- one of the two allelic elements of each of the (1 to N) genotypic indices of either zygotic index of the mating pair node matches either all the first or all the second allelic elements of each of the corresponding (1 to N) genotypic indices of the zygotic index of the zygote node, and
- at least one of the two allelic elements of each of the (1 to N) genotypic indices of the *other* zygotic index of the mating pair node matches either all the first or all the second allelic elements of each of the corresponding (1 to N) genotypic indices of the zygotic

index of the zygote node, whichever allelic elements, the first or the second, were not matched to fulfill condition (i).

These edges must be weighted by coefficients that reflect recombination, meiosis, and, when appropriate, meiotic drive using the general rules for doing so.

6.2.3.4 Haplo-diploid lifecycle graphs. Haplo-diploid models are suitable for hymenoptera, but they also work for so-called "X-linked genes," alleles that are borne by non-hymenoptera on the X-chromosome only, such that females have two copies and males one. Haplo-diploid models are recursions on zygote frequency variables.

Draw $Sbs \times \left[\left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right) + \prod_{k=1}^{N} A_k \right]$ zygote nodes. Assign $Sbs \times \prod_{k=1}^{N} A_k$ nodes square-bracketed type indices containing a N round-bracketed one-place genotypic indices and a sex index of 1. Assign 1/*Sbs* of these each substructure index from 1 to *Sbs* such that no two individual are assigned the same index. There are A_k possibilities for each of the k genotypic indices for these nodes. Assign $1/A_x$ of the nodes each value from 1 to A_x to its xth genotypic place for $x = 1$ to N, such that no two nodes share an index. Assign $Sbs \times \left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right)$ of the initial nodes square-bracketed type indices containing a N round-bracketed two-place genotypic places and sex index 2. Assign 1/*Sbs* of these each substructure index from 1 to *Sbs* such that no two individual are assigned the same index. The type index of each node will consist in a square-bracketed zygotic index containing N genotypic indices featuring for two allelic placesapiece. To generate the type indices for the nodes with sex index 2, combine the alleles at each of the A_k loci to generate variant genotypic indices. Permute these variant genotypic indices with each other to

generate variant zygotic indices. Assign the zygotic indices to the zygote nodes such that no two nodes have the same index.

Next, draw $Sbs \times \left[\left(\prod_{k=1}^{N} \sum_{i=1}^{A} i \right) \times \prod_{k=1}^{N} A_k \right]$ nodes to represent mating pairs. Assign each a sex index of zero. Join each the mating pair nodes to the zygotes nodes using the connect-and-compound rule for sex-dependent selection. The edges just generated must be weighted by assortative mating parameters for whichever zygotes form mating pairs discriminately using the discriminate union of individuals rule.

Once again, the nodes of the second graph are duplicates of the first in the standard fashion. Draw an edge between a mating pair node and a zygote node with a sex index 1 if the allelic element of each *k* of the genotypic indices of the zygote node matches either allelic element of each of the corresponding k^{th} genotypic indices of the mating pair node. Draw an edge between a mating pair node and a zygote node with sex index 2 provided the following conditions hold

- each of the first allelic elements of each of *k* genotypic indices of the zygote node matches each of the allelic element of each of the *k* genotypic indices of the first zygotic index of the mating pair node, and
- each of the second allelic elements at each of the *k* genotypic indices of the zygote node match either of the allelic elements of the corresponding *k* genotypic indices of the second zygotic index of the mating pair node.

These edges must be weighted by coefficients that reflect recombination, meiosis and, when appropriate, meiotic drive, according to the usual rules.

The haplo-diploid models just considered are the last sorts of models I will consider in the graph-drawing phase of the algorithm. It is certainly the case that the above graph-drawing instructions are not comprehensive. There are systems, at least bacteria and polyploids, whose lifecycles cannot be graphed according to the above rules. My hope is that it will be possible, at least in principle, to generate appropriate graph-drawing instructions that will play the same role as do the above ones for systems for which the above graph-drawing rules are inadequate.

6.3 SUBSTITUTION RULES

We now come to substitution portion of the algorithm. The idea in this section is to generalize the equations that can be derived from the lifecycle graphs by weighting the relative frequency terms that appear in the equations by coefficients and functions that quantify the causal influence of ecological, sexual, and individual causal influences. Here we will also deal with subgroup formation.

By putting in values for variables, we can always go from a more determinable model to a version of it, that is, from a model that applies to a wider range of systems to a model that applies to a more narrow range of systems. But we can equally go the *other way* to generate a more determinable model from a more determinate one by weighting variables in the more determinate equation with coefficients. Indeed, we can generalize by weighting with functions as well as coefficients. The equations yielded by the algorithm so far are all determinate versions of equations that are suitable for populations subject to ecological, sexual, and individual causal influences. We generate systems of equations suitable for such populations by introducing coefficients and functions into the equations, ones that generalize the equations by weighting the relative frequency terms in the equations.

The substitution rules we deploy target specific individuals at specific lifecycle stages and hence are deployed over RHS terms of one system of equations at a time. Applying the substitution rules to only one system of equations at a time allows us to recognize that causal influences may function differently at different lifecycle stages. Just because the zygotes in a population are subject to a certain sort of causal influence does not mean that the gametes that formed them were subject to the same causal influence too. (If different individuals in the lifecycle are indeed subject to the same causal influences, then the substitution rules must be deployed repeatedly, once for each lifecycle stage.) Indeed, much of the point of drawing two separate graphs yielding two separate systems of equations was to make it possible to manipulate the equations representing what happens at one lifecycle stage without altering the equations that represent what happens at a previous or a subsequent lifecycle stage.

Here is an example of a system of four equations. The second two equations result from generalizing the first two through the introduction of fitness parameters that weight the relative frequencies of two different types of haploid individual:

$$p'_{1,0,1} = \frac{p_{1,0,1}}{p_{1,0,1} + p_{2,0,1}}$$

$$p'_{2,0,1} = \frac{p_{2,0,1}}{p_{1,0,1} + p_{2,0,1}}$$

$$p'_{1,0,1} = \frac{w_{1,0,1}p_{1,0,1}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1}}$$

$$p'_{2,0,1} = \frac{w_{2,0,1}p_{2,0,1}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1}}$$

Equations 6.10

The remainder of this section will be devoted to stating rules for making the above sorts of generalizations.

6.3.1 Some conventions for stating substitution rules

Here are some conventions I use to state substitution rules:

- the arrow symbol, "→," will be used to separate the term being replaced on the left and the replacing term on the right
- Capital letters represent variables and lower-case letters represent variables at values.
- I use a vertical bar, as in the expression, $X\left|X = x_i\right.$, to indicate a variable at a value. The vertical bars reads "such that." The whole expression reads "variable X such that variable X takes value x."
- (x) will be to mean "for all $X\left|X = x\right.$," (y) will be to mean "for all $Y\left|Y = y\right.$," (z) will be to mean "for all $Z\left|Z = z\right.$."

For example, consider the substitution rule:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z} \rightarrow w_{X|X=x,Y|Y=y,Z|Z=z}p_{X|X=x,Y|Y=y,Z|Z=z}$$

The rule instructs one to replace each relative frequency term in a system of equations that takes definite values for its type, sex, and substructure indices with a product consisting of the same term with the same index values and another variable, a w variable traditionally used for relative fitness, bearing an index that matches that of the relative frequency term with which it is paired. The above rule could be used to effect the transformation of equations (1) and (2) into equations (3) and (4) above.

The annoying use of the vertical bar formalism, along with the "such that" talk, results from the fact that I will need to state two different sorts of rules that I cannot distinguish without

the vertical bar formalism. Indeed, I need to state three sorts of rules that need to be kept distinct. The first sort of rule involves making substitutions on terms that take a specific value for a specific variable. The second sort of rule involves making substitutions on terms that take some value for a specific variable, but can take any value. The last sort of rule involves making substitution on terms by weighting them with expressions involving variables that do not take specific values. For an example of the first sort of thing, consider these equations, a pair of equations appropriate for a haploid population separated into two (hard) substructures:

$$p'_{1,0,1} = \frac{p_{1,0,1}}{p_{1,0,1} + p_{2,0,1} + p_{1,0,2} + p_{2,0,2}}$$

$$p'_{1,0,2} = \frac{p_{1,0,2}}{p_{1,0,1} + p_{2,0,1} + p_{1,0,2} + p_{2,0,2}}$$

$$p'_{2,0,1} = \frac{p_{2,0,1}}{p_{1,0,1} + p_{2,0,1} + p_{1,0,2} + p_{2,0,2}}$$

$$p'_{2,0,2} = \frac{p_{2,0,2}}{p_{1,0,1} + p_{2,0,1} + p_{1,0,2} + p_{2,0,2}}$$

Equations 6.11

It is possible that the ecological causes operating in one substructure are different from the ones operating in the other substructure. If we want to introduce ecological causal influences proper only to the first substructure we need this rule:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=1} \rightarrow w_{X|X=x,Y|Y=y,Z|Z=1}p_{X|X=x,Y|Y=y,Z|Z=1}$$

Note that the *Z* index is fixed at a particular value such that the rule commands weighting the relative frequency terms by fitness values only in the first substructure. Making the prescribed substitutions yields these equations:

$$p'_{1,0,1} = \frac{w_{1,0,1}p_{1,0,1}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1} + p_{1,0,2} + p_{2,0,2}} \quad (7)$$

$$p'_{1,0,2} = \frac{p_{1,0,2}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1} + p_{1,0,2} + p_{2,0,2}} \quad (8)$$

$$p'_{2,0,1} = \frac{w_{2,0,1}p_{2,0,1}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1} + p_{1,0,2} + p_{2,0,2}} \quad (9)$$

$$p'_{2,0,2} = \frac{p_{2,0,2}}{w_{1,0,1}p_{1,0,1} + w_{2,0,1}p_{2,0,1} + p_{1,0,2} + p_{2,0,2}} \quad (10)$$

Equations 6.12

Note that the terms picking out type frequencies in the substructure two are not transformed. The next two sets of rules, the ones for introducing subgroups and for introducing average fitness parameters into a system of equations, involve the use of variables that are not set at particular values.

6.3.2 Subgroup formation rules

The next substitution rules that must be deployed before any of the others can be deployed are ones that represent subgroup formation. Recall from earlier that subgroups are distinct from substructures insofar as substructures are permanent features of the causal landscape. Except for migrants, individuals in a substructure produce descendants in the same substructure; most importantly, they form MICERS only with fellow substructure members. In contrast, subgroups are temporary conglomerations of individuals that form and dissolve within a single lifecycle stage.

Subgroup formation is only interesting insofar as members of different subgroups are exposed to different causal influences, but I recognize the presence of subgroup formation among cotemporaneous individuals before considering how things go differently in each

subgroup. The reason for doing this will be discussed after the rule is stated. But the policy is to form subgroups first and then consider later what goes on differently in each, when substitutions are made for sexual, ecological, and individual causal influences.

Subgroups are always formed from members of the same substructures. We recognize subgroups by making substitutions on the relative frequency terms in a system of equations. The substitution rule for recognizing the existence of subgroups among individuals at a lifecycle stage within a specific substructure is this one (I'll explain the formalism in a moment).

For subgroups formed in some substructure $Z = z$,

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z} \rightarrow \sum_{S} c_{X|X=x,Y|Y=y,Z|Z=z,S}\, p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

For all other individuals,

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z} \rightarrow p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

The second rule simply adds a subgroup index for all individuals that are not found in subgroups. This index is a dummy index; it allows simpler statements of other substitution rules later. The first rule, one for substructures in which subgroups are formed, divides individuals within a specific substructure at a single lifecycle stage into indefinitely many subgroups, each distinguished by taking on a different value for the new index, *S*, which represents in which subgroup the individuals are found. The relative frequency of each type is weighted with a sum of *c* parameters that are also indexed by type, sex, substructure, and subgroup. These are used to pick out the proportion of entities that enter each subgroup. Because they pick out relative frequencies, the *c* parameters associated with one relative frequency term must together sum to 1.

The *c* parameters need not be fixed parameters; they may be set by functions, most plausibly, functions of the relative frequency of types in the model. For instance, Godfrey-Smith and Kerr consider a case in which the *c* parameters are fixed by a binomial sampling function (Godfrey Smith and Kerr 2002, 484), which amounts to imagining that individuals in the model form subgroups with each other indiscriminately.

Subgroups are nested within substructures, so the rule is deployed over the members of *some* particular substructure. In contrast, the rule makes the substitution mandatory for *each* type of individual of *each* sex. Whenever a substitution recognizing subgroups is performed, the *Z* variable must be assigned a value corresponding to the substructure in which the subgroups are found. If subgroups are found in more than one substructure, the substitution rule must be deployed repeatedly for each substructure.

So far, we have not said how selection is going differently in each subgroup. We will do so later when subgroups will be treated as contexts against which to measure the pervasiveness and discrimination of causal influences. This order of operations may strike the reader as odd. Subgroups have been introduced upfront, before consideration of how causal influences operate within them in specific ways that make the subgroups different from each other.

This order was undertaken because the existence of the barriers to causal influence provides grounds for modeling the causal influence of individuals as restricted to individuals in the same subgroup. That subgroups form barriers to causal interaction of exactly the same number and exactly the same size from generation to generation in a recursive model has got to be recognized as a structural feature of the population being modeled. Subgroups are not formed out of individuals that *happen* to causally interact in each generation, because if we formed them in this way, their number and size could be expected to fluctuate from generation to generation.

Rather, individual causal influences go the way they do in each generation because of how the population is structured in the same way in each generation. Maynard Smith (1964) famously imagined haystacks as producing subgroups in which mice could interact; ants provide subgroup contexts for the parasite *Dicrocoelium dendriticum*, notorious for producing brainworms (Sober and Wilson 1998); live rabbits form subgroups in the case of the *mixoma* virus. In each of these cases, there is some persistent fact about the environment faced by individuals in the system that accounts for its subgroup formation, hence it makes sense that it should be possible to recognize subgroup formation without first understanding how selection goes differently in each subgroup.

Lastly, I note that while the subgroup formation rule can be stated using a Riemannian sum, that sum will have to be written out explicitly, without the Riemannian sum operator for later substitution rules to be deployed over the terms making up the sum. Indeed, we will have to perform substitutions on terms that refer to members of particular subgroups in order to say how things go differently in, say, the first subgroup rather than the second.

6.3.3 Hard selection, soft selection and average fitness

In all subgroups, and all substructures, individual causal influences among population members are restricted to individual who are members of the same subgroup or, if there are no subgroups, members of the same substructures. *Soft subgroups* and *soft substructures* pose barriers to competition between the members of distinct groupings, as well as barriers to individual causal influences. What makes soft selection mathematically distinct is the use of an average fitness parameter to which only the members of the grouping make a contribution. Hard subgroups and hard substructures form barriers to individual causal influences, but not competition. What makes hard selection distinct is the use of an average fitness parameter to which all contemporaneous individuals in the population make a contribution.

To see that the barriers to individual causal influences and competition can circumscribe different groupings, imagine a population of rodents separated into two substructures by a river that limits migration between the substructures. The rodents act as individual causal influences over one another's progress, say by engaging in hawk-dove interactions of the sort handled using frequency-dependent selection models. The rodents are also preyed upon by hawks that act to regulate the population size of the rodents. The hawks do not, however, respect the boundary posed by the river. The hawks only eat so many of the rodents and success in avoiding predation by rodents on one side of the river leaves more predators circling to devour rodents on *both* sides of the river. Hence, the rodents compete with others across the river, but do not engage face to face hawk-dove interactions with them. The rodent population is accordingly separated into hard substructures.

I note that "hard" and "soft" have counter-intuitive implications; the barriers between hard subgroups are actually in some sense "softer" than are the barriers between soft subgroups, since the former are not barriers to competition while the latter are. I am simply following the use of language already in play in discussions of variable selection in my use of the hard/soft contrast. Christiansen (1975) adopted the vocabulary of hard and soft from Wallace (Wallace 1968, 1975). It is now clear that Christiansen's use of these notions was different from that of Wallace, who himself has admitted that he was not terribly clear about the contrast he was trying to establish initially. We are, at any rate, stuck with imperfect terminology.

Additionally, I note that subgroups will automatically count as hard subgroups if they last only a fraction of a lifecycle stage, even if barriers between subgroups form barriers to competition, because the individuals who were members of the subgroups compete after having been isolated. In populations that form subgroups, competition between members of different

subgroups occurs if different subgroups can contribute different numbers of members to the next lifecycle stage.

6.3.4 Average fitness substitution rules

Once we have a grip on whether we are dealing with a population with subgroups or substructures that form barriers to competition between cotemporaneous individuals, we can decide what sort of average fitness parameter to deploy for our system. Each relative frequency term, then, must be weighted by the reciprocal of the sum of the frequencies of types against which the individual competes. Note that this means inserting expressions that use variables, not variables-at-values. There are two sets of rules below, one for populations in which competitors causal influences interact with sex differences, that is, they exhibit gene-by-sex interactions, and ones for populations in which they do not.

6.3.4.1 Average fitness rule for soft subgroups and no gene-by-sex interactions. For individuals in soft subgroups, make the following substitutions for each $S = s_j$:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s_j} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s_j}}{\sum_X \sum_Y c_{X,Y,Z|Z=z,S|S=s_j} p_{X,Y,Z|Z=z,S|S=s_j}}$$ Note

that this rule requires a distinct substitution for every type of individual in each subgroup. This sort of substitution is only possible if the Riemannian sums used to express subgroup formation are broken up such that the system of equations includes a distinct term for each individual in each subgroup.

6.3.4.2 Average fitness rule for soft substructures and no gene-by-sex interactions. For individuals in soft substructure $Z = z$, but not in subgroups, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S}}{\sum_{X}\sum_{Y} c_{X,Y,Z|Z=z,S} p_{X,Y,Z|Z=z,S}}$$

6.3.4.3 Average fitness rule for populations without soft subgroups, soft substructures, or gene-by-sex interactions. For individuals neither in subgroups nor soft substructures, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S}}{\sum_{X}\sum_{Y}\sum_{Z} p_{X,Y|Y=y,Z,S}}$$

.Note that we use the above rule for systems of individuals that form hard substructures as well as ones that do not.

6.3.4.4 Average fitness rule for soft subgroups and gene-by-sex interactions. For individuals in soft subgroups, make the following substitutions for each $S = s_j$:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s_j} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s_j}}{\sum_{X} c_{X,Y|Y=y,Z|Z=z,S|S=s_j} p_{X,Y|Y=y,Z|Z=z,S|S=s_j}}$$

Note that this rule requires a distinct substitution for every type of individual in each subgroup. This sort of substitution is only possible if the Riemannian sums used to express subgroup formation are broken up such that the system of equations includes a distinct term for each individual in each subgroup.

6.3.4.5 Average fitness rule for soft substructures and gene-by-sex interactions. For individuals in soft substructure $Z = z$, but not in subgroups, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S}}{\sum_{X} c_{X,Y|Y=y,Z|Z=z,S} p_{X,Y|Y=y,Z|Z=z,S}}$$

6.3.4.6 Average fitness rule for populations without soft subgroups or soft substructures, but with gene-by-sex interactions.

For individuals neither in subgroups nor soft substructures, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow \frac{p_{X|X=x,Y|Y=y,Z|Z=z,S}}{\sum_{X}\sum_{Z} p_{X,Y|Y=y,Z,S}}.$$

6.3.5 Causal influences by substitution

We are now in a position to consider how an array of causal influences over population dynamics can be introduced into population genetics models by substitution. I will consider ecological causal influences first, dividing these into the pervasive and non-pervasive sort. I will then consider individual causal influences. I then consider ecological causal influences. Lastly, I will consider causal influences that have must be modeled using abstract functions, first considering the functions used to model individual causal influences and then considering some specialized models involving density-dependent and temporally variable selection.

6.3.5.1 Pervasive Ecological Causes. Modeling the causal impact of pervasive ecological causal influences is a matter of weighting relative frequency terms to incorporate their impact

individuals' progress through the lifecycle.[28] Recall that we are dealing here with might-as-well-be-ecological causal influences. Pervasive causal influences can be blamed on any pervasive feature of the context of the individual being weighted, as discussed earlier (section 5.2.2.2). It is standard to blame something in the ecological environment. When population geneticists write, in what from the perspective of this work is an informal way, about how relative fitness values represent the impact of "selection" (e.g., Rice 2004, 9), "selection" refers the impact of a pervasive might-as-well-be ecological interactive causal influences. Representing the impact of pervasive ecological causal influences is a matter of simply weighting, with a relative fitness coefficients, the relative frequency terms of individuals that are beset by the causal influence.

Technically, it is non-null pervasive interactive causal influences, ones that have different effects on different types in the model, that require us to introduce weights into the algorithm as discussed in what follows. We must also be careful not to count causes twice when they are part of the same causal chain that influences some types in an interactive fashion. The presence of malarial parasites and the presence of mosquitoes are both causes of the persistence of the sickle-cell allele among humans in Africa, but they should not be counted as distinct such that substitutions are made twice on the same equations to represent each influence, since the mosquitoes provide the mechanism for the spread of parasites.

Pervasive ecological causal influences can be pervasive with respect to the members of a subgroup, a sub-structure, or the population as a whole, so we will need several substitution rules to recognize the various scopes under which pervasiveness can be evaluated.

[28] Strictly speaking, substitutions on every rival type of individual are not absolutely necessary, since we are dealing with *relative* fitness values and we can arbitrarily assign any one type of individual a relative fitness value equal to 1, and then not make a substitution on the relative frequency variable for that type, but I do not work in this shortcut.

6.3.5.1.1 Pervasive ecological causes in subgroups rule. For a pervasive ecological causal influence besetting all individuals in subgroup $S = s$ and substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow w_{X|X=x,Y|Y=y,Z|Z=z,S}p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s}$$

6.3.5.1.2 Pervasive ecological causes in substructures rule. For an ecological causal influence besetting all individuals in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow w_{X|X=x,Y|Y=y,Z|Z=z,S}p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

6.3.5.1.3 Pervasive ecological causes in populations rule. For an ecological causal influence besetting all individuals in a population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow w_{X|X=x,Y|Y=y,Z|Z=z,S}p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

Note that much of the difference between the above rules lies in language of the preamble: the first rule applies to individuals in a specific subgroup in a specific substructure, the second rule applies to all individuals in a substructure, and the last one applies to all population members.

6.3.5.2 Non-pervasive ecological causes. These causal influences require us to split up the members of a grouping into those individuals beset by the causal influence and those beset by its correlatives. Here is where we avail ourselves of *partitions*. Recall that we must always model a non-pervasive causal influence together with its correlatives, even if those correlatives consist in nothing more than the absence of the causal influence. The substitution rules for non-pervasive ecological causal influences are essentially the same as those for pervasive causal influences, expect instead of substituting a fixed relative fitness value into the equation, we substitute a

function into the equation, specifically a Riemannian sum featuring *h* parameters reflecting the rates at which individuals are struck by each correlative of the ecological cause, along with fitness parameters quantifying its impact. All the *h* values sharing type and substructure indices must together sum to 1. The addition of the Riemannian sum function requires adding a fifth index to the relative frequency terms that distinguish the individuals in terms of partition membership.

6.3.5.2.1 Non-pervasive ecological causes in subgroups rule. For an non-pervasive ecological causal influence besetting individuals in subgroup $S = s$ and substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow$$

$$\sum_{k} h_{X|X=x,Y|Y=y,Z|Z=z,S|S=s,k} w_{X|X=x,Y|Y=y,Z|Z=z,S|S=s,k} p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s,k}$$

6.3.5.2.2 Non-pervasive ecological causes in substructures rule. For a non-pervasive ecological causal influence besetting all individuals in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow$$

$$\sum_{k} h_{X|X=x,Y|Y=y,Z|Z=z,S,k} w_{X|X=x,Y|Y=y,Z|Z=z,S,k} p_{X|X=x,Y|Y=y,Z|Z=z,S,k}$$

6.3.5.2.3 Non-pervasive ecological causes in populations rule. For a non-pervasive ecological causal influence besetting all individuals in a population:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow$$

$$\sum_{k} h_{X|X=x,Y|Y=y,Z|Z=z,S,k} w_{X|X=x,Y|Y=y,Z|Z=z,S,k} p_{X|X=x,Y|Y=y,Z|Z=z,S,k}$$

To illustrate the use of the above rules, here is how to use this rule to generate the paradoxical selection on rival diploids with different camouflage discussed in the chapter on the requirements for selection. Beginning at the point at which we have systems of equations for a non-hierarchically structured population with zygote frequencies weighted by an average fitness parameter, we have the following two systems of equations for our system, the first representing the gamete to zygote lifecycle stage and the second representing the zygote to gamete stage (I have dropped the dummy "*S*" index for subgroup membership):

$$p_{[(1)(1)],0,1} = p_{<(1)>,0,1} \bullet p_{<(1)>,0,1}$$
$$p_{[(1)(2)],0,1,1} = p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}$$
$$p_{[(2)(1)],0,1} = p_{<(2)>,0,1} \bullet p_{<(1)>,0,1}$$
$$p_{[(2)(2)],0,1,1} = p_{<(2)>,0,1} \bullet p_{<(2)>,0,1}$$

$$p'_{<(1)>,0,1} = \frac{p_{[(1)(1)],0,1} + p_{[(1)(2)],0,1} + p_{[(2)(1)],0,1}}{p_{[(1)(1)],0,1} p_{[(1)(2)],0,1} + p_{[(2)(1)],0,1} + p_{[(2)(2)],0,1}}$$
$$p'_{<(2)>,0,1} = \frac{p_{[(2)(2)],0,1} + p_{[(1)(2)],0,1} + p_{[(2)(1)],0,1}}{p_{[(1)(1)],0,1} + p_{[(1)(2)],0,1} + p_{[(2)(1)],0,1} + p_{[(2)(2)],0,1}}.$$

Equations 6.13

We weight the zygote frequency parameters in a model of two rival types of diploids according to the above rule for non-pervasive ecological causal influences within an entire population. For our case of two non-pervasive ecological causal influences, we get these equations for the zygote stage of the lifecycle:

$$p'_{<(1)>,0,1,1} = \frac{h_{[(1)(1)],0,1,1} w_{[(1)(1)],0,1,1} p_{[(1)(1)],0,1,1} + h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{[(1)(2)],0,1,1} + h_{[(2)(1)],0,1,1} w_{[(2)(1)],0,1,1} p_{[(2)(1)],0,1,1}}{\overline{w}} +$$

$$\frac{h_{[(1)(1)],0,1,2} w_{[(1)(1)],0,1,2} p_{[(1)(1)],0,1,2} + h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{[(1)(2)],0,1,2} + h_{[(2)(1)],0,1,2} w_{[(2)(1)],0,1,2} p_{[(2)(1)],0,1,2}}{\overline{w}}$$

$$p'_{<(2)>,0,1,1} = \frac{h_{[(2)(2)],0,1,1} w_{[(2)(2)],0,1,1} p_{[(2)(2)],0,1,1} + h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{[(1)(2)],0,1,1} + h_{[(2)(1)],0,1,1} w_{[(2)(1)],0,1,1} p_{[(2)(1)],0,1,1}}{\overline{w}} +$$

$$\frac{h_{[(2)(2)],0,1,2} w_{[(2)(2)],0,1,2} p_{[(2)(2)],0,1,2} + h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{[(1)(2)],0,1,2} + h_{[(2)(1)],0,1,2} w_{[(2)(1)],0,1,2} p_{[(2)(1)],0,1,2}}{\overline{w}}$$

where

$$\overline{w} = h_{[(1)(1)],0,1,1} w_{[(1)(1)],0,1,1} p_{[(1)(1)],0,1,1} + h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{[(1)(2)],0,1,1} + h_{[(2)(1)],0,1,1} w_{[(2)(1)],0,1,1} p_{[(2)(1)],0,1,1} +$$
$$h_{[(2)(2)],0,1,1} w_{[(2)(2)],0,1,1} p_{[(2)(2)],0,1,1} + h_{[(1)(1)],0,1,2} w_{[(1)(1)],0,1,2} p_{[(1)(1)],0,1,2} + h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{[(1)(2)],0,1,2} +$$
$$h_{[(2)(1)],0,1,2} w_{[(2)(1)],0,1,2} p_{[(2)(1)],0,1,2} + h_{[(2)(2)],0,1,2} w_{[(2)(2)],0,1,2} p_{[(2)(2)],0,1,2}$$

Equations 6.14

If we collapse the lifecycle stages, specifying zygote frequencies in terms of gamete frequencies, add the frequencies of the heterozygotes, we end up with this single system of equations, a standard variable selection model:

$$p_{[(1)(1)],0,1} = p_{<(1)>,0,1} \bullet p_{<(1)>,0,1}$$

$$p_{[(1)(2)],0,1,1} = p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}$$

$$p_{[(2)(1)],0,1} = p_{<(2)>,0,1} \bullet p_{<(1)>,0,1}$$

$$p_{[(2)(2)],0,1,1} = p_{<(2)>,0,1} \bullet p_{<(2)>,0,1}$$

$$p'_{<(1)>,0,1,1} = \frac{h_{[(1)(1)],0,1,1} w_{[(1)(1)],0,1,1} p^2_{<(1)>,0,1} + 2h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}}{\overline{w}} +$$

$$\frac{h_{[(1)(1)],0,1,2} w_{[(1)(1)],0,1,2} p^2_{<(1)>,0,1} + 2h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}}{\overline{w}}$$

$$p'_{<(2)>,0,1,1} = \frac{h_{[(2)(2)],0,1,1} w_{[(2)(2)],0,1,1} p^2_{<(2)>,0,1} + 2h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}}{\overline{w}} +$$

$$\frac{h_{[(2)(2)],0,1,2} w_{[(2)(2)],0,1,2} p^2_{<(2)>,0,1} + 2h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1}}{\overline{w}}$$

where

$$\overline{w} = h_{[(1)(1)],0,1,1} w_{[(1)(1)],0,1,1} p^2_{<(1)>,0,1} + 2h_{[(1)(2)],0,1,1} w_{[(1)(2)],0,1,1} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1} + h_{[(1)(1)],0,1,2} w_{[(1)(1)],0,1,2} p^2_{<(1)>,0,1}$$
$$+ 2h_{[(1)(2)],0,1,2} w_{[(1)(2)],0,1,2} p_{<(1)>,0,1} \bullet p_{<(2)>,0,1} + h_{[(2)(2)],0,1,1} w_{[(2)(2)],0,1,1} p^2_{<(2)>,0,1} + h_{[(2)(2)],0,1,2} w_{[(2)(2)],0,1,2} p^2_{<(2)>,0,1}$$

Equations 6.15

If we let every $h = 0.5$, and let $w_{1,0,1,1,1} = w_{2,0,1,1,2} = 1.2$, $w_{[(1)(2)],0,1,1} = w_{[(1)(2)],0,1,2} = 1$, and $w_{1,0,1,1,2} = w_{2,0,1,1,1} = 0.8$, we get a paradoxical variable selection model of the sort discussed as a counterexample to the claim that selection requires heritable traits in chapter 2.

6.3.5.3 The causal influence of sex. The next sort of causal influence that can be handled by substitution rules is the causal influence of sex. Because the individuals in some of our models are indexed by differences in sex, accounting for the causal influence of sex is straightforward, a matter of weighting relative frequency terms in a sex-sensitive way. One can equally think of these causal influences as ecological causal influences that affect different individuals differently by virtue of their sex differences. Recall, too, that gametes have sexes, or at least they have sex indices that reflect their sex of origin, and this makes it possible to consider the causal influence of their having one or another sex of origin in a fashion that parallels how sex differences among zygotes may impact dynamics.

6.3.5.3.1 Sexual causes for males in subgroups rule. To introduce causal influences of being male in subgroup $S|S = s$ in substructure $Z|Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=1,Z|Z=z,S|S=s} \rightarrow p_{X|X=x,Y|Y=1,Z|Z=z,S|S=s}v_{X|X=x,Y|Y=1,Z|Z=z,S|S=s}$$

6.3.5.3.2 Sexual causes for males in substructures rule. To introduce causal influences of being male in substructure $Z|Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=1,Z|Z=z,S} \rightarrow p_{X|X=x,Y|Y=1,Z|Z=z,S}v_{X|X=x,Y|Y=1,Z|Z=z,S}$$

6.3.5.3.3 Sexual causes for males in populations rule. To introduce causal influences of being male in a population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=1,Z|Z=z,S} \rightarrow p_{X|X=x,Y|Y=1,Z|Z=z,S}v_{X|X=x,Y|Y=1,Z|Z=z,S}$$

6.3.5.3.4 Sexual causes for females in subgroups rule. To introduce causal influences of being male in subgroup $S|S=s$ in substructure $Z|Z=z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=2,Z|Z=z,S|S=s} \rightarrow p_{X|X=x,Y|Y=12,Z|Z=z,S|S=s}v_{X|X=x,Y|Y=12,Z|Z=z,S|S=s}$$

6.3.5.3.5 Sexual causes for females in substructures rule. To introduce causal influences of being male in substructure $Z|Z=z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=2,Z|Z=z,S|S=s} \rightarrow p_{X|X=x,Y|Y=2,Z|Z=z,S|S=s}v_{X|X=x,Y|Y=2,Z|Z=z,S|S=s}$$

6.3.5.3.6 Sexual causes for females in populations rule. To introduce causal influences of being male in a population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=2,Z|Z=z,S|S=s} \rightarrow p_{X|X=x,Y|Y=2,Z|Z=z,S|S=s}v_{X|X=x,Y|Y=2,Z|Z=z,S|S=s}$$

Here, the *v* parameters quantify the causal influence of sex on individuals' progress. Note that the only difference between this rule and the rule for handling pervasive ecological causal influences is the use of fixed sex parameters, equal to one for males and two for females.

6.3.5.4 Individual causal influences. The individual causal influences of interest now are ones that are either interactive or discriminate causal influences arising among all the individuals at a lifecycle stage within some grouping, ones that are not MICER-forming causal influences. Because we decided earlier that it was impossible for causal influences to be non-pervasive within some subgrouping without being pervasive within some smaller grouping, we need only

consider pervasive individual causal influences. Hence we need only a two rules to handle individual causal influences.

The central component of the strategy for dealing with causal influences emanating from individuals is to depend upon causally interpretable frequency-dependent relative fitness *functions* to capture the causal influences of individuals on one another. Frequency-dependent selection models provide a vehicle for modeling individual causal influences because the causal influence of individuals at a lifecycle stage are contingent upon their relative frequency. The extent of the impact of one type of individual on the progress of others will be modulated by how many individuals of that type there are in the population.

What we want is a rule that states how we should weight the relative frequencies of types within groupings to reflect the causal influences that all the types upon each others progress. But here we encounter the obstacle that individuals may causally influence one another in any number of different ways. One option, then, is simply to substitute a general function into the equation to represent individual causal influences. I use $f_{X,S}(p)$ as a function taking the (possibly weighted) absolute numbers of all the types of individual in the model as arguments. So the influence of individuals over each others' progress could be represented as follows.

6.3.5.4.1 Individual causal influences in subgroups rule. For a individual causal influences among individuals in subgroup $S = s$ in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow f_{X,S}(p)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s}$$

6.3.5.4.2 Individual causal influences in substructures rule. For a individual causal influences among individuals in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(p)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

6.3.5.4.3 Individual causal influences in populations rule. For a individual causal influences among individuals in an entire population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(p)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

As stated, however, the substitution rules cannot be used as part of a set of rules that yield an equation with which inferences about population dynamics can be made, because nothing about the nature of the function relating relative frequencies to progress has been stated. All the above substitution rules do is signal that the relative frequency terms within subgroups or substructures must be weighted by *some* function of the relative frequency of the types in the subgroups or substructures; the substitution rule does not say which function to use. In principle, any function could be used here.

The function one most commonly finds in textbooks discussions of frequency-dependent selection is an additive function (Hedrick 2005, 223). Such a function is appropriate for *non-interactive* causal influences stemming from population members, and because this sort of function is so common, it makes sense to write down what frequency-dependent selection looks like from this viewpoint. For individual interactions among the members of an entire population, that function would set

$$f_{X,S}(p) = \sum_{X} \beta_X \alpha_X p_{X|X=x,Y|Y=y,Z|Z=z,S} .$$

The alpha parameters in the above rule quantify the causal influence of the types they weight on the type whose relative frequency is being replaced. They can take any value, including a value of 1 for individuals that exert no influence on their fellows. The β parameters represent the extent to which such causal influences are *discriminate*, and the β parameters that share a subscript must sum to 1. Discriminate individual causal influences will be appropriate for cases in which

individuals, for instance altruists, can recognize one another as such, perhaps by their green beards, and direct their altruism accordingly.[29]

6.3.5.5 Temporally variable selection and density-dependent selection. There are two more kinds of population genetics models that involve substituting functions into basic equations, temporally variable selection models and density-dependent selection models. In each case, the causal influence of types of individuals in the model is contingent upon some parameter, either time or population size. Individual causal influences presented us with our first taste of what is probably the most serious problem with the algorithm offered here, namely, that it cannot help much with models in which relative frequency variables are weighted by functions that take as arguments variables that appear elsewhere in the model. That problem is especially pronounced in the cases of temporally variable selection and density-dependent selection. There is no one general way in which population size might causally influence the progress of a kind of individual, and there is no one general way in which variable environments must vary over time. Temporally varying interactive ecological causal influences are especially intractable: they may include indefinitely many correlatives, the correlatives may replace one another often or rarely, ecological causal influences may vary in a random fashion, or occur in a sequence, and so on.

It is interesting to note that, though temporally variable selection models and spatially variable selection models are usually presented together and their polymorphism-yielding properties often compared, from the current perspective, they are very different in terms of how easily they yield to a general analysis. Spatially variable selection models, though complex, yield to analysis much more easily, and this seems largely because of the availability of the Riemannian sum operator, which allows one to state general rules for handling non-pervasive ecological causal influences, no matter how many of these there are, no matter how strongly their

[29] For a case of altruism of this sort, see Keller and Ross (1998).

impact is correlated with the variant types of individuals in the model, and no matter how many individuals are beset by each sort in the each generation. "Spatial" differences in the influence of ecological causes can be presented all at once and left to recur in each generation. By contrast, little can be said, in a general way, about temporally variable selection models because there is no single general way to represent different times all at once in a fixed *recursive* equation.

For what it is worth, temporally variable selection models at least get off the ground when one replaces relative frequency terms with ones paired with functions of time. Letting *t* range over different times, we can at least write the following sorts of substitution for individuals within subgroups and substructures respectively, where *t* is time in generations:

6.3.5.5.1 Temporally variable causal influences in subgroups rule. For a temporally variable causal influences among individuals in subgroup $S = s$ in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow f_{X,S}(t)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s}$$

6.3.5.5.2 Temporally variable causal influences in substructures rule. For a temporally variable causal influences among individuals in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(t)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

6.3.5.5.3 Temporally variable causal influences in populations rule. For a temporally variable causal influences among individuals in an entire population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(t)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

For density-dependent selection, we not only substitute in functions, this time functions of population size, but also substitute in census population size parameters for all the relative frequency terms. All classical population genetics models are in fact mathematical simplifications of models featuring absolute numbers of variant types of individuals (see Rice 2004, 10). Because each type is a member of the same population, and most population genetics models feature additive average fitness parameters in their denominators, the census population size parameter cancels out of every numerator and every denominator term in all models except those featuring density-dependent selection. The approach pursued here is to keep things simple and deal with relative frequency terms whenever possible, when absolute numbers of individual are needed, we couple each relative frequency term with a census population size variable.

We must insert census population size variables not only for RHS variables but, unusually, for LHS variables too, since we are explicitly representing the consequences of changing population size in a population in which the progress of individuals is tied to the changing overall size of the population. Thus we must make two sets of substitutions to countenance density-dependent selection, we must first replace every RHS and LHS individual relative frequency term by a term representing the absolute number of individuals:

6.3.5.5.4 Population size introduction rules. For populations in which population size is a causal influence, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow Nc \bullet p_{X|_{X=x},Y|_{Y=y},Z|_{Z=z},S|_{S=s}}$$

$$(x)(y)(z)p'_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow Nc \bullet p'_{X|_{X=x},Y|_{Y=y},Z|_{Z=z},S|_{S=s}}$$

By weighting these relative frequency terms by census population size, we are creating terms that refer to absolute numbers of individuals. Then we must make a general substitution

representing the causal influence of population size by weighting relative frequency terms by functions of *Nc*.

6.3.5.5.5 Density-dependent selection in subgroups rule. For a density dependent causal influences among individuals in subgroup $S = s$ in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s} \rightarrow f_{X,S}(Nc)p_{X|X=x,Y|Y=y,Z|Z=z,S|S=s}$$

6.3.5.5.6 Density-dependent selection in substructures rule. For a density dependent causal influences among individuals in substructure $Z = z$, make the following substitutions:

$$(x)(y)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(Nc)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

6.3.5.5.7 Density-dependent selection in populations rule. For a density dependent causal influences among individuals in an entire population, make the following substitutions:

$$(x)(y)(z)p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow f_{X,S}(Nc)p_{X|X=x,Y|Y=y,Z|Z=z,S}$$

Once again, such equations are useless until some determinate function for $f_{X,S}$ (*Nc*) has been specified, but once again, anything is possible here. Roughgarden (1971) considers several possibilities. Of course, causal knowledge of the impact of population size will supply the function in specific cases, but the algorithm does not do so, and probably cannot do so, in the general case.

6.3.5.6 Mutation. The last substitution rule we deploy is one that injects into our equations terms that reflect mutation rates, the rate at which individuals of one type produce individuals of another type. Here is the substitution rule for the introduction of mutation parameters:

$$(i)\, p_{X|X=x,Y|Y=y,Z|Z=z,S} \rightarrow p_{X|X=x,Y|Y=y,Z|Z=z,S} + \sum_{X} u_{X,X|X=x} p_{X,Y|Y=y,Z|Z=z,S}$$

To see the rule in operation, take a simple model featuring three haploid individuals not separated into substructures, for which the following system of equations would emerge from the graphing stage of the algorithm:

$$p'_{1,0,1} = p_{1,0,1}$$
$$p'_{2,0,1} = p_{2,0,1}$$
$$p'_{3,0,1} = p_{3,0,1}$$

Equations 6.16

To add mutation, we make these substitutions:

$$p_{1,0,1} \rightarrow p_{1,0,1} + \sum_{X=1}^{3} u_{X,1} p_{X,0,1}$$

$$p_{2,0,1} \rightarrow p_{2,0,1} + \sum_{X=1}^{3} u_{X,2} p_{X,0,1}$$

$$p_{3,0,1} \rightarrow p_{3,0,1} + \sum_{X=1}^{3} u_{X,3} p_{X,0,1}$$

Note how the sum features a variable, *X*, picking out each rival type. The result of applying the last set of substitution rules is the following system of equations:

$$p_{1,0,1} = p_{1,0,1} + \sum_{X=1}^{3} u_{X,1} p_{X,0,1}$$

$$p_{2,0,1} = p_{2,0,1} + \sum_{X=1}^{3} u_{X,2} p_{X,0,1}$$

$$p_{3,0,1} = p_{3,0,1} + \sum_{X=1}^{3} u_{X,3} p_{X,0,1}$$

Equations 6.17

If we expand the sums, we get these equations:

$$p_{1,0,1} = p_{1,0,1} + u_{1,1} p_{1,0,1} + u_{2,1} p_{2,0,1} + u_{3,1} p_{1,0,1}$$

$$p_{2,0,1} = p_{2,0,1} + u_{1,2} p_{1,0,1} + u_{2,2} p_{2,0,1} + u_{3,2} p_{1,0,1}$$

$$p_{3,0,1} = p_{3,0,1} + u_{1,3} p_{1,0,1} + u_{2,3} p_{2,0,1} + u_{3,3} p_{1,0,1}$$

Equations 6.18

Mutation parameters with identical subscripts represent the rates at which individuals of the type picked out by the subscript produce individuals of other types and hence must be negative. Each of the *u* parameters with non-identical subscripts reflects the rate at which individuals whose type is picked out by the first subscript on the *u* are produced by individuals whose type if picked out by the second subscript. Thus, $u_{1,1}$ reflects the rates at which type one individuals produce descendants of other types, while $u_{1,3}$ represents the rate at which individuals of type 1 are produced by individuals of type 3. The mutation parameters must add up such that a parameter with identical subscripts is the sum of all the others sharing its first subscript.

6.4 EQUATIONAL COLLAPSE

We have now reached the point at which we have equations that represent all the causal information acting at each stage of the lifecycle that can be represented in textbook discrete generation Wright-Fisher models and can be generated by the algorithm. The last thing we need to do is substitute the right-hand sides of the equations generated by the first graph into the RHS equations for the second graph, for systems featuring two graphs. We do this by replacing variables on the right-hand side of the second system of equations with the entire right-hand side of equations from the first system that specify them according to the first set of equations. See section 6.2.3.3.2.2 for an example of this sort of collapse.

7.0. DRIFT

In this section, I will briefly consider the notion of drift in selection theory. More specifically, I am interested in what causal influences are responsible for the stochastic element of selection theory, an element that is quantified, in Wright-Fisher models of non-neutral evolution, by the parameter, *variance effective population size* (*Nev*). My stance is that this parameter quantifies the causal influence of NINPICs, non-interactive, non-pervasive, indiscriminate causal influences.

My aim is to keep this section short. Those interested in a more elaborate defense of my claim that the notion of drift, when used in its causal sense, refers to NINPICs may consult a paper I have written on the subject (Gildenhuys 2008). That paper also includes criticisms of rival approaches to understanding "drift" in its causal sense, specifically those of the pairs of Millstein and Beatty (Beatty 1984, 1992; Millstein 1996, 2002, 2005), and Rosenberg and Bouchard (Rosenberg and Bouchard 2004; Rosenberg and Bouchard 2005). But the reader will recall that NINPICs, along with pervasive non-interactive causes, are leftovers from the deterministic cause-to-model algorithm.

Pervasive non-interactive causes have at best an indirect causal influence over population dynamics; they are the sorts of things that could affect census population size and, by extension, *Nev*. But pervasive non-interactive causes could not possibly swing population dynamics in one way or another directly because they affect all population members in the same way. Indeed, pervasive non-interactive causes are hard even to imagine, for they would have to alter the

descendant production of all population members to same extent, and it is hard to imagine a natural phenomenon affecting a population in so precise a manner. That leaves NINPICs as the only thing left that could be responsible the stochastic element of population genetics models.

Before expanding that last argument that the impact of NINPICs are what is quantified by *Nev*, I will offer a brief discussion of the causal notion of drift, and how that notion functions in Wright-Fisher models and the diffusion approximation to those. After that, I discuss how NINPICs control the value taken on by a population for *Nev*.

7.1 THE CAUSAL NOTION OF DRIFT

My goal is to explicate the causal dimension of drift, to say what is causally responsible for the stochastic element of population genetics models, and hence to say what is responsible for the randomness of the dynamics of the populations that these models govern. Population geneticists often use "drift" in a causal sense, especially when characterizing diffusion theory. Hedrick discusses how "the relative impact of genetic drift and selection varies with the population size" (Hedrick 2005, 352). Gillespie characterizes drift as a "source of randomness" (Gillespie 1998, 82). Robertson, in the famous paper in which he develops the notion of the retardation factor, writes that "gene frequency changes under the combined effects of selection and drift due to finite population size are determined to a good approximation by *Ns*" (Robertson 1962, 222). In those sorts of usages, "drift" is being used to refer to a certain sort of cause, something that has effects, and something that can be laid alongside selection such that its influence can be combined with it and the extent of its influence contrasted with it.

In Wright-Fisher and diffusion theory models involving selection and drift operating simultaneously on population dynamics, the influence of drift is quantified by the variance

effective population size parameter (*Nev*). We will be able to explicate the causal notion of drift when we can say what sorts of causes are in play in populations that have the sort of influence that "drift" is supposed to have in these models, causes whose influence is quantified by *Nev*. Indeed, we are better off thinking of the link between NINPICs and *Nev* as direct, a link unmediated by the notion of drift: once we have understood what causes *Nev* represents in population genetics models, we have grasped what we need to understand in order to understand the deployment of stochastic models in selection theory, so that it no longer matters what we call drift. Given the ambiguity of the term, we should probably not refer to *anything* as drift.

It is perhaps worth quickly noting why I regard conceptions of drift one finds in population genetics texts as inadequate for the project of understanding what is quantified by *Nev*. Rice calls selection, mutation, and migration "directional factors" and contrasts these with drift, which he calls a "non-directional process" and sometimes a "non-directional effect" (Rice 2004, 131-132; 135). The non-directionality of drift manifests in its tendency to cause the spread of any of the rival types within a population with equal likelihood. Thus, drift leads to random changes in allele frequency (Gillespie 1998, 19; Jacquard 1974, 164; Gale 1990, 13). Hartl and Clark put the point in terms of predetermination: "Because the sampling process does not change the allele frequencies in a predetermined way, this process is known as random genetic drift" (2007, 95).

Instead of picking out drift in terms of the sort of influence that it has, as a non-directional sort of influence, a non-predetermined sort of influence, or a random change-producing influence, we should instead seek a formulation of drift that defines these non-directional factors in other terms. A good definition of drift will pick out the non-directional factors without reference to their impact on population dynamics and accordingly can be used as

part of an *explanatory* account of why population dynamics are stochastic. If we want to explain why population dynamics involve a non-directional component by appeal to the causal influence of drift, we cannot do so if we define drift as what produces the non-directionality.

7.2 VARIANCE EFFECTIVE POPULATION SIZE IN POPULATION GENETICS

We use the variance effective population size parameter as the population size parameter in the binomial sampling equation in Wright-Fisher models; we also use it to set values for the variance term in diffusion theory, a continuous approximation available for some Wright-Fisher models. This last deployment is actually the main deployment of the *Nev* in classical population genetics models, because under the diffusion approximation, one can draw analytic results for the dynamics of populations in which NINPICs have not been idealized away. Without the diffusion approximation, the derivation of analytic results for Wright-Fisher models is not possible.

For example, the forward Kolgomorov equation in diffusion theory characterizes population dynamics by way of a two term equation:

$$\frac{\partial \Psi(p,t)}{\partial t} = -\frac{\partial}{\partial p}[\Psi(p,t)M(p)] + \frac{1}{2}\frac{\partial^2}{\partial p^2}[\Psi(p,t)V(p)]$$

Equation 7.1

with the psi function representing the time-dependent probability distribution function for the relative frequency p of a target allele. The first term on the right, featuring the M function, governs directional processes, quantifying the tendency for the probability distribution to flow in the direction in which selection and other directional processes (mutation, migration) are working. The value of the M function is set using deterministic Wright-fisher population genetics equations involving fixed relative fitness values, mutation rates, and migration rates. The second term on the right is the variance term featuring the V function, which governs how the density of

the probability distribution tends to flow into adjacent regions, both to the left and to the right. The variance term involves "non-directional influences" insofar as it governs the rate of flow of probability from one allele frequency to both higher and lower frequencies.

That the variance term is a function of the second derivative of the probability density function reflects the fact that net flow into a region of the distribution is positive whenever the change in the slope of the distribution at that region is positive, such that there is more net flow into the region than there is out of it. So if the probability density to the right (say) of a given allele frequency is a little bit smaller than the probability density at the chosen allele frequency, while the density to the left is much larger than the density at the chosen allele frequency, probability will "flow" into the region from the left. Over time, the non-directional influence captured by the variance term will cause the distribution to spread out (Rice 2004, 132); because of drift, the probability flows to unoccupied regions, and to regions where the differences in the differentials between it and surrounding regions is especially large. The greater the variance term, the greater the probability of finding the allele frequency of the modeled population further from where it was initially likeliest to be found.

When using diffusion theory as a continuous approximation to the discrete generations model of Wright and Fisher, the variance term is set by equation for the binomial sampling variance (Rice 2004, 140):

$$V = \frac{p(1-p)}{2Nev}$$

Equation 7.2

Thus, the smaller the value for *Nev* taken on by a population, the larger is its variance term. Whatever causes contribute to the value taken on by *Nev* will thus control the extent to which

population dynamics involves the sort of non-directional dynamics captured by the variance term in diffusion theory.

In discrete generation models, the probability of that an allele goes from one relative frequency to another is represented using a probability transition matrix, one with an entry for every possible change in allele frequency. From the matrix, one can read off the probability that a population will go from any allele frequency to any other allele frequency. The probabilities in the probability transition matrix are set, for the two allele diploid case, using the binomial sampling equation:

$$x_{ij} = \frac{(p')^{2Nev-i}(1-p')^{i}(2Nev)!}{(2Nev-i)!i!}$$

Equation 7.3

where *Nev* is the variance effective population size, *p'* is the adult (post-selection) allele frequency as calculated deterministically, $i = p' \cdot 2Nev$ and $j = (1 - p') \cdot 2Nev$ (see Hedrick 2005, 351). When the variance effective population size is large, the probabilities in the transition matrix representing small changes in relative frequency are larger, while those representing bigger jumps in the relative frequency are smaller. As *Nev* gets smaller, the probability of larger changes in relative frequencies between generations grows. Thus, here again, we see the effective population size controlling the extent to which population dynamics are stochastic, that is, the extent to which population are likely to evolve further from their current relative frequencies.

I said earlier that it is only after a deterministic model has been deployed that we can add in the stochastic element to Wright-Fisher models. The above equation shows why. The p' term, which is fed into the binomial sampling equation, is generated using a deterministic model. So it is only after we have deployed a deterministic model that we are in a position to deploy the

equation that is appropriate for population dynamics that include a stochastic element. It is worth noting, too, that only Wright-Fisher models that are recursions on allele frequencies can be fed into the above binomial sampling equation. Furthermore, because the equation is a binomial sampling equation, the approach is appropriate only for systems consisting of two rival alleles, or two rival competitors; the notion of *Nev* is not defined for multi-allelic systems (Ewens 2004, 127). I do not know how to quantify the impact of NINPICs on populations whose population dynamics cannot be captured using models that are recursions on two alleles.

7.3 NON-INTERACTIVE, NON-PERVASIVE, INDISCRIMINATE CAUSAL INFLUENCES MAKE POPULATION DYNAMICS STOCHASTIC

The argument that NINPICs are what is being referred to as "drift" when that notion is being used in a causal sense is simply that nothing else will do. This should not be surprising. Besides NINPICs and pervasive non-interactive causes, all other causal influences on population dynamics are taken into account in the cause-to-model algorithm. Since pervasive non-interactive causal influences could not possibility be responsible for the stochasticity in population genetics models since they affect all population members in the same way, that leaves NINPICs as the only sort of cause remaining that could account for the stochastic character of population genetics models.

Each of the criteria for sorting causes that I deploy, interactivity, discrimination, and pervasiveness, is such that any cause must count as either one or the other. All causal influences are either interactive or non-interactive. All causal influences are either discriminate or indiscriminate. All causal influences are either pervasive or non-pervasive. However, any cause that counts as either interactive, or pervasive, or discriminate, or any combination of these,

simply cannot exert a "non-directional" influence over population dynamics. Recall that exerting a non-directional influence over population dynamics is our criterion of adequacy for an explication of the causal dimension of the notion of drift.

Consider interaction first. *Interactive* causes of reproduction among population members have different causal influences on the reproduction of the different types in the population. Causes with this sort must be handled through the introduction of relative fitness parameters into the model. By influencing the reproduction of population members in a type-sensitive fashion, as interactive causes must do, they induce the reproduction of one type and impede that of another, at least when these are measured relative to one another. This sort of influence will be either directional or stabilizing; either way, we can eliminate interactive causes from those that might count as drift.

Non-interactive causal influences on reproduction that are *pervasive* cannot induce drift. These are just the causes left aside along with NINPICs when the deterministic algorithm for generating population genetics models was considered. Pervasive causes have no potential to influence *which* population members form a sample of reproducers from a larger sample of would-be reproducers, and hence cannot have a stochastic influence over population dynamics.

Non-interactive, *discriminate* causes cannot produce non-directional evolution either. If a non-interactive causal influence affects reproduction, and differentially affects one type of population member, such a cause will either promote or inhibit the spread of the type with which it is statistically associated, depending on whether or not it is beneficial. This sort of causal influence produces a directional influence, and hence fails to have the sorts of effects that drift is supposed to have.

7.4 NON-INTERACTIVE, NON-PERVASIVE, INDISCRIMINATE CAUSAL INFLUENCES AS A MODULATING CAUSAL INFLUENCE

NINPICs are the influences that Rice calls "non-directional"; they explain the stochasticity in population genetics. However, I want to stress that while I have characterized *Nev* as representing the causal influence of NINPICs on population dynamics, *Nev* is nonetheless sensitive to other causal influences too. Discriminate causal influences, such as those produced by linkage of target alleles to other alleles undergoing selection, impact the effective size of the target allele populations (Santiago and Caballero 1998). Furthermore, population size and fluctuations therein provide mechanisms by which *any* sort of causal influence can impact *Nev*, and not all of these influences must be NINPICs. And of course many demographic features of a population impact *Nev* too. Inbreeding produces a situation in which alleles of the same sort tend to be destroyed by NINPICs together; sex ratio bias impacts *Nev* too. I point these things out so that it is clear that my claim that NINPICs explain the stochasticity in population genetics is *not* based on the stance that NINPICs, and only NINPICs, impact the value taken by a population for *Nev*.

My thesis is this: the fact that NINPICs beset a population *makes it make sense* to treat the dynamics of that population as a stochastic affair. I do not claim that there are not alternative treatments of natural populations, real or imaginary, that do not involve probabilities. I claim only that when we deploy selection theory, what licenses our deployment of the sorts of probabilities we deploy as part of the theory is that populations are invariably beset by NINPICs.

The special importance of NINPICs in selection theory can be brought out in the following way: while *Nev* is sensitive to all the causal influences of the variables that turn up in its derivations, the presence of NINPICs is a necessary condition for those parameters to matter

to *Nev* in the first place. So while there are lots of causal influences on *Nev*, NINPICs are special among these because their presence is a precondition for the others to have any effect at all, and it is in that sense that *Nev* "represents" the impact of NINPICs. All other causes of stochasticity in population genetics are at best *modulating* causes of stochasticity.

To see this, imagine a population without NINPICs but with every last variable with an impact on effective population size set anywhere you please. Imagine any level of inbreeding, any sort of population structure, any variance in offspring number, etc…. Still the dynamics of such a population will not be a matter of chance. Conversely, again set all the causal influences on *Nev* anywhere you please and add NINPICs into the picture; stochasticity of the sort typical of evolutionary theory will result. Hence my (somewhat awkwardly versed) claim that NINPICs are what *make it make sense* to model population dynamics as stochastic in evolutionary theory: without NINPICs we have no reason to deploy an effective population size parameter in evolutionary theory at all.

8.0 CONCLUSION

In the conclusion, I reconsider the question of just how general is the approach to selection theory taken here. This is something I promised to do earlier. I divide the problem of generality into two components. I first consider the generality of the algorithm I offer for generating dynamical models for those systems that fit my entrance rule. I then consider the question of whether the entrance rule for selection theory that I have proposed might itself be too narrow in scope.

Before getting to those discussions, however, I want to re-iterate a couple of claims made in the introduction. The first is that while my use of causal vocabulary to present selection theory generalizes the theory (though perhaps not fully or as much as it should), the use of that vocabulary accomplishes more than just generalization. Even more important than generalization is how the use of causal vocabulary to present the theory exposes why selection theory should be regarded as having an explanatory structure. The second claim I want to repeat from the first chapter is that even if it should turn out that what I carve out as selection theory is an arbitrarily delimited subsection of a broader field of causal models, what would then be the not-quite-general-enough generalization accomplished above would not be completely undermined. Rather, it would be consist in a partial accomplishment of project that could be furthered.

8.1 THE GENERALITY OF THE ALGORITHM

The algorithm offered in the previous chapter has three limitations that bear immediate discussion. One is that I effectively assume that some causal influences do not interact; another is that the algorithm fails to shed light on the character of some functions that appear in it, specifically ones that take as arguments variables that appear in the equations outside the scope of the function; a final limitation is that it makes heavy use of the language of genetics.

8.1.1 The multiplicative collapse of fitness assumption

By stating the rules for handling individual and ecological causal influences in the fashion that I have, I have implicitly been taking advantage of an assumption I dub the multiplicative collapse of fitness assumption, the assumption that sexual, individual, and ecological causes operative in a population do not interact with other such influences. For example, two ecological causal influences do not interact if an individual beset by one ecological causal influence will not react to it any differently if it beset by one or another correlative of a non-pervasive ecological cause.

To see how the assumption that causes introduced by substitution do not interact is undertaken, note that the substitution rules are deployed without regard to what other causal influences have already been taken into account in the population. For instance, if two pervasive ecological causal influences beset the population, one simply deploys the substitution rule for pervasive ecological factors twice, once for each. This means that one simply ends up forming a product from the coefficients reflecting each ecological causal influence. Hedrick advocates this sort of procedure when considering a population of plants that face different environmental pressures at different times in their development from germination to maturity. In Hedrick's model, the coefficients representing the impact of each episode of selection are simply multiplied together in the mathematical model for the system (Hedrick 2005, 177). He calls the assumption

implicit in the procedure the *multiplicative collapse of viability*; it is from Hedrick that I got the name for my more general assumption, the multiplicative collapse of fitness.

To see how the assumption is limiting, consider how ecological causal influences might well interact. For instance, if subjection to some correlative of a non-pervasive ecological causal influence E1 is harmful, and equally subjection to another one E2 is also harmful, but subjection to both is advantageous, it will not be possible to countenance the impact of E1 and E2 by weighting relative frequency parameters by a single coefficient for each.

Could I discharge the assumption of the multiplicative collapse of fitness? I suspect that one could, but I will not try to do so here. For one thing, the assumption of multiplicative fitness collapse is widespread when applied to ecological causes. The fact that a population genetics textbook deploys a narrower version of the multiplicative collapse of fitness assumption does redeem it a little. Nonetheless, I note my undertaking of the assumptions of the multiplicative collapse of fitness as a real failure.

If one did try to discharge this assumption, things would get quite complicated when it came to non-pervasive causes that interacted with others or even just co-varied with them. In such cases, one would have to get a grip on the extent to which subjection to one of several possible correlatives of one cause was statistically associated with subjection to each of several correlatives of another non-pervasive cause, where each different correlative of one causal influence could potentially interact with each distinct correlative of the other. If one followed the practice of partitioning the population into distinct causal contexts, what is perhaps the natural way to handle such interactions, the number of partitions would increase very rapidly as more causes and correlatives of causes were added to the picture. Accordingly, an algorithm that permitted consideration of interactions among the sorts of causes I introduce by substitution

would have to be extremely complicated unless a general rule could be developed that reduced the handling of all interactions among the causes to a single policy that could be stated in a general way, without respect to the number of interacting causes. I have not so far been able to think up such a rule.

8.1.2 Unspecified Functions

Another limitation of the algorithm I offer is that it cannot be used to specify the character of a definite subset of functions that appear in classical population genetics equations. The algorithm is unhelpful about the character of function that take as arguments variables that appear elsewhere in the equations for the dynamics of the system, though sometimes only implicitly. As I noted earlier, frequency-dependent selection, temporally variable selection, and density-dependent selection all involve modeling the impact of some quantity that is at least implicitly already in play in the basic equations used to model selection in simpler cases: relative frequency variables are what are weighted by the frequency-dependent selection functions used to model the impact of individuals on their contemporaneous fellows; time is implicitly in play in the contrast between next-generation frequencies and this generation frequencies (e.g., p' vs. p); and population size must be determined before countenancing any causal influences over population dynamics (chapter 4). Moreover, various parameters in population genetics models are often functions of relative frequency variables, even though the algorithm does not present them as such. As a couple examples of this last sort of thing, consider assortative mating parameters and homing parameters.

Consider the case of assortative mating first. In the algorithm as I have set it up, assortative mating functions are fixed parameters. But the disposition of one type of zygote to form a mating pair with another type of zygote must be sensitive to the relative frequency of the

various types of zygotes in the population, and this sensitivity may manifest itself in a variety of ways. Invariant dispositions to pair with the especially attractive types of zygote may be unrealistic for populations in which there may not be enough of the attractive zygotes to go around. We might expect other zygotes to be more disposed to settle on a less strongly preferred mate when the frequency of the preferred mate is low, but to be choosier when there are many of the preferred types around. The solution to this sort of difficulty is to present assortative mating parameters as complex functions of relative frequency parameters in the model, as is done for instance in Seger (1985).

The algorithm does not say what function to use to set assortative mating parameters, and different ones may be appropriate for different populations. The algorithm fails by not being informative enough: it does not tell you what features of populations to look for to assign which functions for assortative mating parameters.

As a second example of this phenomenon, consider the case of values for the *h* parameters from the non-pervasive ecological causal influence rule used to quantify the dispositions of individuals to be subjected to correlative of a non-pervasive ecological cause rather than another. The *h* parameters can be understood as homing parameters; they specify the extent to which individuals of each generation manage to seek out niches in which they are subject to a specific variety of non-pervasive causal influence in the environment, most plausibly ones especially conducive to their development. Sometimes, homing parameters are simply fixed coefficients, but homing parameters may also be presented as functions of relative frequencies, too. Again, realism is the motivating factor, as there may not be sufficient room in a preferred niche when the relative frequency of the variant that prefers that niche gets sufficiently high. Several theorists have suggested a variety of functions for homing parameters to accommodate

this (Templeton and Rothman 1981; Garcia-Dorado 1986; Hedrick 1990). Once again, the algorithm I developed does not present parameters measuring the extent to which an ecological influence is discriminate as functions of relative frequency variables. I present them as fixed parameters, even though they ought to be specified by functions in at least some cases.

To be sure, there is nothing in the algorithm that *prevents* anyone elaborating it by replacing parameters with functions. The algorithm is not incompatible with more elaborate approaches to population genetics models that use functions in the place of parameters. Still, the failure of the algorithm fails to shed any light on when and how parameters should be specified by functions is a real limitation.

8.1.3 The language of genetics

We now come to the last of the major limitations of the algorithm: it makes heavy use of concepts proper to genetic systems. The reader might accordingly be suspicious of that usage in the context of this work, which is supposed to be a general account of selection theory, one that is supposed to present the theory as applicable to systems of genetic variations. Before embarking on the algorithm, I traded in the official vocabulary of this work, cause-talk, for the standard vocabulary of population genetics (allele, genotype, gamete, zygote, mating pair, etc). My ostensible reason for doing so was that the standard vocabulary was familiar and that I had offered parallel concepts using cause-talk for the vocabulary of population genetics (competitor, MICER, etc…). Because such things as gametes, zygotes, and mating pairs had been declared instances of more general notions understood in causal terms, the deployment of traditional vocabulary in the place of causal notions was deemed harmless and convenient.

Still, the clear reliance on specifically genetic vocabulary raises the concern that the algorithm might not carry the same generality as is sought in this work because it will not work

for some competitors that are not alleles. Insofar as the algorithm is known to work, it is known to work because it generates textbooks models for the dynamics of systems delineated by their genetics variations. But perhaps there are competitors that are not enough like alleles for the algorithm to work for them, but are nonetheless still competitors according to my definition, and hence cannot be correctly processed by the algorithm. This difficulty is serious, because it brings into question whether there really is such a thing as selection theory, that is, whether *selection theory* is something other than just *classical population genetics*.

One initial way that I can defend myself from the accusation that my algorithm has a more limited scope than my entrance rule is simply to point out that I have already been explicit that the algorithm has a more narrow application than does the entrance rule. The algorithm applies only to systems with discrete generations, or at most to ones that can be treated *as though* they had discrete generations. It also does not work for polyploids or bacteria that engage in parasexual processes. The reader can review section 5.3 where I make explicit all the limitations to the scope of the algorithm that I could recognize, though there may be more than I countenance there. It should be recognized that population geneticists deploy age-structured models as well as a considerable variety of alternative approaches to modeling selection, including Moran and Cannings models, the statistical moment approach, and approaches that take advantage of the formalism of quantitative genetics, each with their unique arrays of guiding assumptions. While all the systems covered by these alternative approaches are *populations of competitors* as I have defined "competitor" and "population," it would not necessarily be appropriate to treat them using the algorithm I present. So my aim was never complete generality anyhow.

The force of the criticism that my entrance rule has a more general application than my cause-to-model algorithm can be further mitigated by pointing out that those who are pushing the importance of non-genetic inheritance systems have done little to show how instances of expanded inheritance could be mathematically modeled. Advocates of expanded inheritance, such as Griffiths and Gray, deploy faulty statements of the requirements for selection and then do little more than point out that such things as chromatin marking schemes may meet these requirements. Until it is demonstrated by the advocates of a particular sort of expanded inheritance *how* one might model these sorts of things, I will be unable to determine whether or not they fit the above algorithm at all. Still, for what it is worth and only as far as I can tell rival variant cell templates and rival host imprints could be treated as rival haploids are in the algorithm, while rival chromatin marking schemes could be treated as are rival alleles among diploids.

However, even though my best guess is that the above algorithm would not have to be adjusted to handle these last cases of expanded inheritance, there is no reason to expect that the algorithm should generally be adequate for handling non-genetic systems. However, I think that demanding that the algorithm apply *as is* to novel inheritance structures, ones not yet considered by able-minded model-generating population geneticists, is to place too heavy a demand upon it. A weaker, and I think more reasonable demand, would be to require that the rules of the algorithm be sufficiently flexible that they could be *adjusted* to take into consideration novel sorts of systems, ones that are not, anyway, alleles or the sorts of individuals that I earlier defined in terms of the relationships borne by their constituent alleles.

Few scientific theories are born full-fledged from the minds of their authors. Generally, they are expanded and amplified by a number of authors over a considerable length of time;

classical population genetics has certainly evolved in this way. My algorithm is supposed to be a formalization of (a fragment of a generalization of) classical population genetics, and, as classical population genetics evolves, the algorithm should accordingly be expected to develop with time, too. At least, it must be the sort of thing that *could* develop with time.

Consider, for example, the hypothetical phenomenon of a species with three sexes. Clearly, the algorithm does not work for such systems since the sex index on the relative frequency terms ranges only from 0 to 2. However, it is fairly clear how the algorithm would have to be adjusted were we to find species with three sexes. We would permit the sex index to take on value 3, we would draw 1.5 times as many nodes representing zygotes as we would draw for otherwise similar systems with two sexes, we would make new rules for forming mating pairs from three-sex genotypes where three edges from nodes with distinct sex indices would converge on nodes representing next generation zygotes, and so on. The algorithm would not completely fall apart in the face of populations with three sexes, but rather it would have to be adjusted to accommodate the novelty.

Similarly, it may turn out that colonial organisms, such as Portuguese man-o'wars, are MICERs formed from several different lineages of organisms. They may well be similar to mating pairs, but with more members. I don't know enough about colonial organisms to tell whether this is the case. But if it did turn out that this was the case, I would have to write a new set of rules for drawing lifecycle graphs for such populations, ones in which more than two edges from the zygote stage would converge on a single node representing the colonial organism formed from zygotes individuated by their genotypes. Other adjustments would undoubtedly have to be made too, depending on the details of the reproduction of the colonial organism. But again, the algorithm would not fall apart but would simply have to be generalized somewhat.

So I claim that if the algorithm's three main components, the decision tree, the graph-drawing rules, and the substitution rules could be adjusted to comprise the novel systems, then the algorithm I present should not be impugned as a means of presenting the inner workings of selection theory, even though it has a more determinate application than does my entrance rule. Even though I have in no way provided a comprehensive account of selection theory, one that shows how to model every last system of competitors of interest, it may remain the case that I have offered a *set of techniques* for presenting such modeling practice in formal terms. In doing so, I have put forward a suggestion for a language that can be used to understand selection-theoretic modeling. Just as one gets a grip on the natural numbers by learning how to count, rather than by counting them all, one can similarly get a grip on systems covered by selection theory through learning a set of techniques such as drawing lifecycle graphs, generating equations, and making substitutions on these to generate sytstems of equations for them.

So selection theory might thus be something more determinate than just the causal modeling of a subset of worldly systems that ultimately bear nothing more in common than merely having been traditionally thought interesting by researchers able to generate dynamical models of population dynamics using implicit causal modeling techniques. At the same time, it might not be something so definite as to be presentable using a finite algorithm. Instead, selection theory might contain the resources to generate indefinitely many variant dynamical models, but require the deployment of a finite set of modeling techniques to do so, ones that can be used repeatedly and in varying combinations. The three main modeling techniques deployed in the algorithm, the decision tree, the lifecycle graphs, and the substitution rules, might provide a vocabulary sufficient to say what is peculiar about every last system that meets the entrance rule for the theory, without providing the means to list them all. If that is the case, then the fact that the algorithm functions only for a fragment of systems that fit the entrance rule is not damning.

8.2 THE GENERALITY OF THE ENTRANCE RULE

In the previous section, the concern was that systems that fit the proposed entrance rule might not be accurately processed by the algorithm I put forward in chapter 6. I now turn to discuss whether the entrance rule itself is too narrow in scope. That discussion will be inconclusive, though hopefully not uninformative. Mostly, it consists in a discussion of how one might approach the problem of determining whether the entrance rule proposed here is sufficiently general. It will turn out that I lack the necessary resources to take anything more than a tentative stance on the question. The discussion leads into a consideration of a number of future projects of research.

In the first chapter of this work, I presented what I have been calling selection theory as a generalization of population genetics. There, I noted that just about everyone who studies natural selection in the abstract undertakes the same stance that I do, that the theory of natural selection applies to more systems than just those considered in its textbook formulations. I take a particular approach to generalizing selection theory: I use causal vocabulary to generalize the individual bits of biological vocabulary one finds in classical population genetics.

It is natural to wonder just how general is the generalization pursued here. In particular, the reader might want to know whether the entrance rule I propose encompasses every last system that has been said to undergo natural selection. If the circumstances under which selection theory may be deployed are set using my entrance rule, does the theory become sufficiently general to capture the dynamics of immunological systems, the dynamics of cultural variations, the development of neurological systems, and other systems that have been said by others to undergo selection (e.g., Hull, Glenn, and Langman 2001; Heinrich, Boyd, and Richerson 2008)?

One possibility is that the generalization pursued above does in fact encompass every system that undergoes selection. Since that possibility is not threatening, let us move on to consider the possibility that the generalization of selection theory pursued in this work only covers some of the systems that undergo selection. The first thing to realize is that this is not the sort of determination at which I can arrive by measuring my entrance rule for selection theory against some other rule for determining what systems undergo selection, because if I had such an alternative rule, I would have used it as my entrance rule for selection theory. However, even without such a rule, we can still pursue the threatening possibility that the above account of selection theory fails to be as general as it could or should be. We can do this provided we have a grip on what would constitute a legitimate generalization of selection theory as I present it.

For the sake of having some definite examples of systems that a legitimate generalization of selection theory might encompass, but which are excluded by the entrance rule that I have proposed, let's consider two sorts of systems said by other writers to be undergoing selection:

- Heinrich and Boyd (2002) consider a system of cultural variants, differing mental representations of the moon. The variants blend into one another on a scale between two extremes. At the first extreme is a conception of the moon as lacking in intentional attributes, while at the opposite extreme is a conception of the Moon according to which it has a full array of folk psychological states. In the model, mental representations tend to produce more extreme descendants, rather than producing descendants of the same type as themselves.
- Szathmary (1999) suggests that replicators from the early history of life on Earth did not compete, though they underwent selection. Because these early replicators impeded their own replication, their dynamics were not marked by the struggle for existence, but instead

> by the "survival of everybody," a situation conducive to the maintenance of a variety of different types of replicators and hence to the possibility of assembling these into more complex systems.

My entrance rule for selection theory definitely excludes both these sorts of systems because I require that entities that undergo selection produce descendants of the same type as themselves and I require that they compete. The moon concepts in Heinrich and Boyd's model do not produce descendants of the same type as themselves; Szathmary's early replicators do not compete.[30]

As part of his (1999) work, Szathmary takes an explicit general stance on the requirements for selection:

1. *Multiplication*. Entities should give rise to more entities of the same kind.
2. *Heredity*. Like begets like; A-type entities produce A-type entities; B-type entities produce B-type entities; etc.
3. *Variability*. Heredity is not exact; occasionally A type objects give rise to A' type objects (it may be that A' = B). (1999, 31)
4.

Szathmary's statement of the requirements for selection and mine are similar. Szathmary's first two conditions are essentially the same as my condition that competitors produce descendants of the same type as themselves. Szathmary gets variation into his system through his third condition, something that I do as part of my competition requirement. Since a system with competitors is a determinate sort of system with variation, my statement of the requirements for selection is more specific than Szathmary's and accordingly his stands as a generalization of mine.

[30] I note that Boyd and Richerson, along with Szathmary, mean to use their models for the same purposes that I pick out as the purposes of selection theory. Heinrich and Boyd's model shows how adaptive evolution of culture toward improved conceptions could occur; Szathmary means to show how a widespread polymorphism among early replicators could occur. So it is not the case that these theorists are pursuing a different type of explanation than that provided by selection theory as I construe it.

If we are to consider whether my account of selection theory is too narrow for excluding the systems said by Heinrich and Boyd and Szathmary to be undergoing selection, we must take it that theories are definite sorts of things. Our criterion for what constitutes a legitimate generalization of a theory must involve some definite position on what are and are not scientific theories. What's more, it will require a more determinate stance on this question than the ones taken up by advocates of the syntactic and semantic views of scientific theories. None of this is obvious, so I argue for these claims now.

To begin, consider how any theory combined with any other theory would yield something more general than either one alone. For instance, selection theory plus quantum mechanics is more general than just selection theory alone. But a generalization of selection theory that combined it with quantum mechanics would produce a strange amalgam; a generalization of this sort would not constitute an advance. Though I have not specified where the distinction lies, the contrast between legitimate generalizations and illegitimate amalgams does make sense: We would all agree that Newton's classical mechanics is a generalization of Kepler's theory of celestial mechanics and a major scientific advance. In contrast, Newton's theory combined with nineteenth century political economy might be a generalization of both theories in some logical sense, but it would not be an advance on either theory.

Since there is no question that selection theory plus quantum mechanics is more general than selection theory, in a strictly logical sense of "general," and further since we must delegitimize amalgams of this sort in some way or another in order even to pursue the question of whether the entrance rule for selection theory offered here is too narrow, therefore we must ban such amalgams on the grounds that they are not scientific theories, which plainly they are not. Thus, negotiating the question of whether one has a sufficiently general account of a theory

requires a definite conception of what constitutes a theory, that is, a rule to determine what does and does not count as a scientific theory, a rule that would capture the intuitively compelling fact that selection theory plus quantum mechanics is not a single scientific theory, and neither is classical mechanics plus political economy. Furthermore, the rule that supplies us with those judgments must present scientific theories as more definite sorts of things than just collections of propositions, families of mathematical models, or systems of partially interpreted equations, since our mere amalgams could easily qualify as any of these.

I do not have a rule for deciding what does and what does not count as a scientific theory, but I think that such a rule is worth pursuing. For one thing, it is worth pursuing because it would help settle the issue of when a generalization of a theory constitutes an advance. Furthermore, it is worth pursuing because the problem of confirmation is driven to no small extent by conceptions of scientific theories that are too permissive. Philosophers have attempted to make explicit under what conditions evidence favors some theory, especially how it favors one theory over another, but many proposals for understanding the relation between theory and evidence have been subject to counterexamples that depend on the legitimacy of conceiving of theories as collections of propositions to which one can add propositions at will. The infamous tacking problem, for instance, depends on our ability to append an arbitrary proposition to a scientific theory for which we have some evidence so as to illegitimately garner evidential support for the tacked on claim. If we are to solve the problem of confirmation, to say what the relationship between some theory and some evidence must be in order for the latter to confirm the former, I suspect that we need to conceive of theories as more definite sorts of things than just collections of propositions to do so. Constraining what counts as a scientific theory would, anyway, provide a simple way to avoid the tacking problem and many similar problems.

Even though I lack a general rule for deciding what counts as a scientific theory, I do have an inferentialist account of one theory, a generalization of classical population genetics. If we assume that this sort of account of what scientific theories are is at least somewhat generalizable, we can consider the question of what a generalization of selection theory would have to look like in order to constitute an advance.

For starters, it should be clear that one could not tack on the claim, "the sky is blue," to selection theory as presented here because this statement cannot be integrated into the set of inferential rules that constitutes the theory. The statement makes no contribution to determining over what sorts of systems the theory can be deployed, it does not help us attach specific natural systems to specific mathematical models, and it is not a bit of mathematics that allows us to calculate system dynamics.

Generalizing from this example, I have some ideas about what a generalization of selection theory should like look like on the inferentialist approach taken here. A generalization of selection theory would constitute a generalization of the rules that constitute the theory. To avoid licensing strange amalgams, I suggest that such generalizations must proceed in a piecemeal fashion, with the rules being generalized one at a time while those that are not being generalized are held fixed. So, a generalization of the entrance rule for the theory would be legitimate if the dynamics of the novel systems allowed in could be inferred using the existing set of rules for attaching natural systems to mathematical models. At least some of the systems that developmental system theorists discuss when advocating for their generalization of selection theory are of this sort: variant lineages of endosymbionts among aphids could be treated as are variant alleles in selection theory (Griffiths and Gray 2001, 198). Generalizations of the rules for assigning dynamical models to systems that already meet the entrance rule for the theory would

equally constitute legitimate advances, provided they apply to the old sorts of systems. All the theoretical advances in population genetics modeling developed after the theory was initially put forward, such as Levene's developmental of variable selection models (Levene 1953), are generalizations of this sort, since they apply to systems of individuals differentiated by genetic variations.

Note that if we don't hold either the entrance rule or the guts of the theory fixed when we generalize the other, we cannot exclude the amalgam of selection theory and quantum mechanics from counting as a legitimate generalization of either theory. To see this, assume for the sake of argument that quantum mechanics can be understood in the same sort of inferentialist terms I have used to understand selection theory. The selection theory/quantum mechanics amalgam would just be a theory with a more general entrance rule and a theory with a greater number of rules for inferring system dynamics. The entrance rule would require that a target system either be a competitor or a system of subatomic particles. The former would be treated with the rules offered above and the latter would be treated with quantum mechanics. The very first rule in the body of the theory would determine whether the system should be treated using the machinery of selection theory or that of quantum mechanics. That machinery would then operate just as before, allowing the scientist to draw implications about system dynamics.

However, if we refuse to allow the guts of the selection theory and the entrance rule for it to be generalized at once, we can prohibit the selection theory/quantum mechanics amalgam. This is so, anyway, if we add a further proviso against useless rules. To see this, note that the dynamics of competitors cannot be calculated in the same fashion as are the dynamics of subatomic particles, so we could not generalize selection theory by adding to it the calculational machinery of quantum mechanics while holding fixed the entrance rule of selection theory. Such

a generalization would simply amount to the addition of useless rules. Similarly, one cannot make truth-preserving inferences about the dynamics of subatomic particles using the calculational machinery of selection theory, so one could not hold that machinery fixed and generalize the entrance rule of selection theory to include subatomic systems.

A slightly weaker proposal than the one just rehearsed would allow the generalization of the entrance rule for the theory coupled with a generalization of *some* the rules for inferring system dynamics, provided that some of the old rules applied to the new sorts of systems. In the algorithm of chapter 6, different systems are treated differently in systematically different ways. We get different lifecycle graphs for diploids and haploids, for instance, ones from which different equations are derived, equations that are then fed into the rules for making generalizing substitutions. Those substitution rules, however, are the same for both diploids and haploids. Even though these two sorts of systems are treated differently initially, they are treated in the same way in other respects later on. A more lenient approach to generalization would allow the introduction of novel sorts of systems along with novel calculation machinery for assigning mathematical models to systems, provided the new calculation machinery could be integrated with old machinery in this sort of way. Wholly new machinery to go with wholly new sorts of systems would still be banned.

I recognize that such talk of the integration of new rules with old ones is indeterminate; I have merely pointed to a sort of integration exhibited by the account of selection theory I have offered to get the idea across. But the selection theory / quantum mechanics amalgam would not, presumably, show this sort of integration. This is because as soon as one had determined that one had a system that fit the entrance rule of the amalgam, one would have decide whether to calculate its dynamics using the rules drawn from selection theory or those drawn from quantum

mechanics. The theory would send each sort of system down a separate stream, where the rules for calculating the dynamics of systems on one stream would no appear among the rules for calculating the dynamic of systems on the other stream.

It is worth noting that the arguments used in the second chapter of this work against rival conceptions of the requirements for selection fit with the understanding of generalization presented here. Besides arguing that circular and incomplete presentations of a theory are faulty, I claimed in chapter 2 that resemblance selectionist accounts of the requirements for selection were too narrow. They excluded systems that should be included as ones undergoing selection. My reasoning in that first chapter was essentially that we can treat the excluded systems using the same apparatus that is used for systems already recognized as undergoing selection. For instance, Lewontin's exclusion of systems that do not evolve, such as ones that exhibit overdominance, is a mistake because the dynamics of such systems can be handled with the same machinery as those that do evolve. Whether or not a system exhibits overdominance is just a matter of the values for the variables in the mathematical model of its dynamics, and the same goes for many other systems that exhibit stable polymorphisms. The upshot is that we can hold fixed the calculational apparatus of the theory, what we use to assign mathematical models to systems, and painlessly allow in systems exhibiting overdominance and ones that do not evolve. We can do the same with systems that do not exhibit the inheritance of fitness or phenotype owing to gene-by-sex, or gene-by-environment interactions.

Assuming that we can in fact negotiate the threatening possibility that the account of selection theory offered here is too narrow in the manner I have suggested, then we have arrived, finally, at a definite way of framing that issue. I earlier asked whether every last system said to undergo selection fit my entrance rule for selection theory. Since no one is ever in a position to

measure a rule they endorse for determining the scope of some theory against another non-identical rule that they think *really* determines the correct scope of a theory, the threatening possibility that the account of selection theory offered here is too narrow should instead be arbitrated by determining whether the account could be generalized in a non-trivial way that generates an advance. So it is now at least possible at least to rigorously pose the question of whether Szathmary's systems of early replicators, and Boyd and Richerson's moon model, are models of selection: can one generalize selection theory to accommodate these sorts of systems so as to yield an advance?

I have not done nearly enough research into the mathematical modeling of early life on Earth to determine whether models of the dynamics of early replicators are generalizations of selection theory according to the criterion just constructed. As noted earlier, Szathmary's statement of the requirements for selection is a generalization of mine, but the question of whether it is a generalization *that is also an advance* hinges on whether the same set of rules for calculating the dynamics of the more determinate set of systems I consider can be deployed over systems that I rule out, but that Szathmary lets in. On the face of it, there is a real possibility that my rules cannot work for the systems that Szathmary, but not I, would regard as undergoing selection, because my rules deploy the notion of competition; most saliently my population circumscription rule requires that candidate population members compete with individuals already part of the population. But perhaps these rules and others in the "guts" of my proposal, ones that feature the notion of competition, could be generalized, too, while other rules from the guts of the theory that do not involve competition could be preserved. This would be legitimate on the more lenient proposal considered above for what constitutes a generalization that is an advance. Here, anyway, is an opportunity for further post-dissertation research.

What about the Heinrich and Boyd moon model? My suspicions are that this model should not be regarded as part of selection theory. My basis for this view is that those authors use the Price equation to calculate the long-run dynamics of their system. Those authors set values for variables in the Price equation with functions that are supposed to specify the causal impacts of the causes in their model, specifically, the attractiveness of unambiguous moon conceptions and the selection pressure against the conception of the moon as having intentionality. This means that the system Heinrich and Boyd consider is very different from those over which selection theory is deployed, because one cannot calculate the dynamics of systems of competitors in this way. Except for a very few systems in which evolution is constant, the Price equation cannot be used to explain the dynamics of systems of variant competing alleles. Indeed, the Price equation cannot even be used even to *infer* the dynamics of systems of competitors, for it requires one to already understand the dynamics of one's system in order to fix the values for its variables.[31]

Thus, insofar as the dynamics of a system can be modeled using the Price equation, the system is sufficiently different from the sorts of populations considered in classical population genetics that it should be treated as failing under the purview of a different theory. A version of selection theory that was general enough to encompass both classical population genetics and the Heinrich and Boyd moon model, with its use of the Price equation as a dynamical modeling tool, would be an amalgam. Such an amalgam would require one to decide right off the bat whether one was dealing with a cultural system of the sort Heinrich and Boyd consider, for which the use

[31] There is an approach to modeling systems of the sort treated in classical population genetics using roughly the same scheme as the Price equation, in which dynamics are the product of differential selection and transmission bias. The approach is exceedingly complex; indeed a two-allele, two-locus system features 255 free coefficients in the function used to compute the impact of differential selection (Kirkpatrick, Johnson, and Barton 2002, 1732). Anyhow, the Price equation, a simple sum of two statistical parameters, is not up to the task of inferring system dynamics.

of the Price equation as a recursive model would be legitimate[32], or a system of variant alleles (or competitors) for which it would (nearly always) fail to be legitimate. Each sorts of system would then be handled each according to its own set of rules for generating mathematical models of system dynamics. If I am right about all this, then, as it turns out, the dynamics of cultural variants, at ones of the sort Heinrich and Boyd consider, are handled using a different theory than are the dynamics of genetic variants.

8.3 FUTURE RESEARCH

At this point, I list and discuss some post-dissertation research projects that grow out of the work done here.

- Many writers the theory of natural selection through consideration of the Price equation and it is one of my post-dissertation projects to write a critique of the widespread contemporary handling of causal and explanatory questions about population dynamics through the formalism provided by the Price equation. This would give my work some relevance to the broader philosophy of biology community, members of which tend increasingly to see the study of selection through the lens of the Price equation. Importantly, the causal construal of what explains population dynamics offered here is at odds with what authors working with the Price equation claim explains the dynamics of the same systems. If the causal explanation of the dynamics of systems that I advocate is correct, the explanations pursued by those who

[32] Presuming it would, in fact, be legitimate. While I am suspicious of Heinrich and Boyd's use of the Price equation as a causal model, I cannot contest that usage in the same way I could context the use of the Price equation as a causal model for the dynamics of systems of alleles. In this latter case, I can show that the Price equation cannot be used to make inferences about long-run system dynamics when its variables are set by fixed functions. This can be done by showing that treating systems of alleles in this way would yield inferences that are incompatible with the ones generated by classical population genetics equations. But I have no alternative formalism for inferring the dynamics of the sort of cultural system that Boyd and Richerson consider.

deploy the Price equation must be wrong. My job talk brought this out with respect to a single model, the most commonly discussed model of "group selection." Talk of group selection as a causal influence over population dynamics quantified by a covariance parameter strikes me as grossly misplaced; my algorithm does not contemplate such causal influences yet it generates "group selection" models. Covariances and other statistical quantities that are labeled "group selection," "collective-level selection," "particle-level selection" may arise owing to a variety of different causal scenarios, they can occur for a variety of different sorts of groupings, and the explanation of dynamics in each case should be different. Furthermore, the Price equation cannot function as a tool for making inferences about the long-run dynamics of natural populations because its right-hand side quantities change in value as the population evolves, and they do so in a way that cannot be inferred using the Price equation. This fact has largely gone unnoticed, though Kerr and Godfrey-Smith recognize it (Kerr and Godfrey Smith 2008, 533). Anyhow, if inference is understood as a necessary condition for explanation, the Price equation cannot be understood as even a candidate explainer of the dynamics of systems governed by classical population genetics.

- I would like to pursue a definite conception of what scientific theories are that is up to two tasks: first, it must make sense of our intuitive judgments about which sorts of generalizations of scientific theories are advances and which are not; second, it must prohibit amalgams of the sort that are offered as counterexamples to accounts of the relationship between theory and evidence. The semantic and syntactic accounts of scientific theories currently on offer are too loose to perform these tasks, and it is not clear to me what work they are supposed to do. The project of delineating scientific

theories could be pursued as part of a more general attempt to contribute to our understanding of how theory relates to evidence, though that problem looks exceedingly difficult to solve. I suspect categorial grammar may be useful as a formalism to approach that last problem though.

- I could initially approach the development of a more general account of scientific theories through consideration of other specific theories. One suspicion I have developed from this dissertation work is that causal notions may lie behind crucial bits of vocabulary and modeling procedures in other special sciences. A natural choice as a test case for this possibility is micro-economics owing to the influence of Malthus over both Darwin and Wallace. Does classical microeconomics have an entrance rule? Are the sorts of systems over which that theory is deployable describable in causal terms? Can an account of which equations get deployed over which systems be generated out of causal understanding? Is a firm a specific sort of causal system? Perhaps even an inferentialist theory of theories is worth pursuing if it turns out that microeconomics and other theories can be tackled in the same way as selection theory has been tackled here.
- I am especially interested in the potential to use my NINPICs to explain error terms in microeconomics and perhaps elsewhere, too. Error terms, u_i, are often treated as random variables with definite probability distributions, in favorable cases these ones:
    - $E(u_i) = 0$,
    - homoscedasticity or constancy of the variance of the u_i across different values of the other RHS variables
    - statistical independence or absence of auto-correlations in the u_i

- statistical independence of the error term and other RHS variables. (Woodward 2003, 316)

The above list is a list of facts about the sort of influence that error terms can be expected to have. It is akin to the definition of drift that makes it a "non-directional" influence. A better definition of what produces error would define error terms by reference to the sorts of causal influences that have the features listed above, that is, those that have the above list of features as implications. I suspect that error terms that have the above four features quantify the influence of NINPICs.

- As it currently stands, the algorithm is in need of work. Here are some more significant improvements I could make:
  - I could capture a greater number of temporally variable selection models if I drew graphs for multiple lifecycles, thereby making it possible to ascribe different relative fitness coefficients to different individuals at different times. As it stands, temporally variable selection models contain unspecified functions, but these could be replaced in at least some instances for systems with multi-generational lifecycle graphs.
  - I could add a consideration Kimura's concept of quasi-linkage equilibrium (Kimura 1965), or perhaps a generalization of it of the sort considered by Kirkpatrick, Johnson, and Barton for use in the statistical moment approach to modeling system dynamics (2002). A population is in quasi-linkage equilibrium when the statistical associations between genes at different loci are evolving slowly enough that they can be treated as effectively unchanging. Under such circumstances, it is possible to deploy single-locus models for populations of

alleles even when alleles at other loci exert discriminate and/or interactive causal influences over them. Under quasi-linkage equilibrium, one *can* average over the influence of alleles at other loci because the extent and character of their influence remains pretty much invariant from generation to generation. In the decision-tree portion of the algorithm, I claimed that a multi-locus model is necessary whenever genes at other loci exert discriminate or interactive causal influences over population dynamics, but in fact that assertion is too strong. Getting a grip on when quasi-linkage equilibrium holds in causal language would put me in a position to modify the decision-tree portion of the algorithm such that it would license a broader deployment of single-locus models, a boon since these are the simplest ones.

- The algorithm could very likely also be modified in more fundamental ways so as to be less unwieldy. Ideally, the algorithm could be written such that either the graphical stage or the substitutional stage of the algorithm is eliminated. Perhaps I could find a way of accomplishing what is accomplished now by the graphical portion of the algorithm using substitution rules. I have not tried as hard as I might to accomplish this, and it may very well be possible. An algorithm that involved only fewer steps would be a tidier affair, even if a large number substitution rules had to be constructed. On the other hand, I could accomplish what is now accomplished by substitution rules through the use of rules for both drawing and *manipulating* causal graphs to produce a figure that shows and quantifies every causal fact necessary to model population dynamics. I did not take this approach in the dissertation because doing so would require deploying

- Lastly, I would like to pursue research into the major transitions in evolution. This would allow me to determine whether the entrance rule for selection theory could be generalized in some way that would allow in replicators that do not compete, as discussed above. But furthermore, one promising avenue of investigation would be to pursue the implications of the distinction between sub-groups and MICERs. Sometimes, the question of how multi-cellularity might have evolved is pursued by treating composite bodies, such as multi-cellular organisms, as *groups* of more basic components, such as single-celled organisms (e.g., Michod 1999). It is a critical feature of population genetics modeling that zygotes and sub-groups are different sorts of things. These two sorts of groupings play mathematically distinct roles in population genetics, and, accordingly, showing how a population might evolve such that its members form subgroups of one sort or another falls short of showing how it might

evolve such that its members form MICERs. So, one avenue for investigation would be to consider whether the sorts of composite bodies considered in the major transitions literature might all be better understood as MICERs than as subgroups, and, if so, whether the problem of how these things could have evolved could be advanced by relocating the focus from models involving subgroup formation to ones involving MICER formation.

BIBLIOGRAPHY

Andrewartha, H.G., and L.C. Birch (1984), *The Ecological Web*. Chicago: University of Chicago Press.

Ariew, Andre, and R. C. Lewontin (2004), "The Confusions of Fitness", *British Journal for the Philosophy of Science* 55 (2):347-363.

Arnold, Anthony J., and K. Fristrup (1982), "The Theory of Evolution by Natural Selection: A Hierarchical Expansion", *Paleobiology* 8 (2):113-129.

Avital, E., and Eva Jablonka (2000), *Animal Traditions: Behavioral Inheritance in Evolution*. Cambridge: Cambridge University Press.

Bateson, Patrick (2006), "The Nest's Tale. A reply to Richard Dawkins", *Biology and Philosophy* 21:533-558.

Bayliss, Christopher D., Dawn Field, and Richard Moxon (2001), "The Simple Sequence Contingency Loci of *Haemophilus Influenzae* and *Neisseria Meningitidis*", *Journal of Clinical Investigation* 107 (6):657-662.

Beatty, John (1984), "Chance and Natural Selection", *Philosophy of Science* 51:183-211.

——— (1992), "Random Drift", in E.K. Keller and Elisabeth Lloyd (eds.), *Keywords in Evolutionary Biology*, Cambridge, MA: Harvard University Press, 273-281.

Boyd, R., and Joseph Heinrich (2002), "Why Cultural Evolution does not Require the Replication of Representations", *Journal of Cognition and Culture* 2 (2):87-112.

Brandom, Robert (1998), *Making It Explicit*. Cambridge: Cambridge University Press.

Brandon, Robert (1990), *Adaptation and Environment*. Princeton: Princeton University Press.

Changeux, Jean-Pierre, and Stanislas Dehaene (1989), "Neuronal Models of Cognitive Functions", *Cognition* 33:63-109.

Christiansen, F. B., and T. Prout (2000), "Aspects of Fitness", in Rama S. Singh and Costas B. Krimbas (eds.), *Evolutionary Genetics: From Molecules to Morphology*, Cambridge: Cambridge University Press.

Christiansen, Freddy Bugge (1975), "Hard and Soft Selection in a Subdivided Population", *American Naturalist* 109 (965):11-16.

Damuth, John (1985), "Selection among Species: A Formulation in Terms of Natural Functional Units", *Evolution* 39 (5):1132-1146.

Darwin, Charles (1988[1859]), *On the Origin of Species: A Facsimile of the First Edition*. Cambridge, MA: Harvard University Press.

——— (1989[1884]), *The Different Forms of Flowers and Plants*. Edited by P.H. Barrett and R.B. Freeman, *The Works of Charles Darwin*. New York: New York University Press.

Dawkins, Richard (1982), *The Extended Phenotype: the Gene as the Unit of Selection*. New York: Oxford University Press.

——— (2004), "Extended Phenotype - But not *Too* Extended. A Reply to Laland, Turner, and Jablonka", *Biology and Philosophy* 19:377-396.

de Meeûs, T., and J. Goudet (2000), "Adaptive Diversity in Heterogeneous Environments for Populations Regulated by a Mixture of Soft and Hard Selection", *Evolutionary Ecology Research* 2:981-995.

Dessalles, Jean-Louis (2007), *Why We Talk: The evolutionary origins of language*. Translated by James Grieve. Cambridge: Oxford University Press.

Dobzhansky, Th. (1970), *Genetics of the Evolutionary Process*. New York: Columbia University Press.

Earman, John, and John Roberts (1999), "*Ceteris Paribus* There Are No Provisos", *Synthese* 118:438-479.

Earman, John, John Roberts, and Sheldon Smith (2002), "Ceteris Paribus Lost", *Erkenntnis* 57 (3):281-301.

Endler, John (2000), "Adaptive Genetic Variation in the Wild", in Timothy Mousseau, Barry Sinervo and John Endler (eds.), *Adaptive Genetic Variation in the Wild*, New York: Oxford University Press.

Ewens, Walter J. (2004), *Mathematical Population Genetics: 1. Theoretical introduction*. Second ed. New York: Springer-Verlag.

Fisher, R.A. (1918), "The Correlation Between Relatives on the Supposition of Mendelian Inheritance", *Royal Society of Edinburgh* 52:399-433.

——— (1930), *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Franklin, I., and R. C. Lewontin (1970), "Is the Gene the Unit of Selection?" *Genetics* 65:707-734.

Futuyama, D. J. (1998), *Evolutionary Biology*. Sunderland, MA: Sinauer Associates.

Gale, Jeff (1990), *Theoretical Population Genetics*. London: Unwin Hyman Ltd.

Gannett, Lisa (2003), "Making Populations: Bounding Genes in Space and in Time", *Philosophy of Science* 70:989-1001.

Garcia-Dorado (1986), "The Effect of Niche Preference on Polymorphism Protection in a Heterogenous Environment", *Evolution* 40:936-945.

Gavrilets, Sergey, and William R. Rice (2006), "Genetic models of homosexuality: generating testable predictions", *Proceedings of the Royal Society B* 273:3031-3038.

Ghiselin, Michael T. (1974), *The Economy of Nature and the Evolution of Sex*. Berkeley: University of California Press.

Gigord, L. D. B., M.R. Macnair, and A. Smithson (2001), "Negative Frequency-Dependent Selection Maintains a Dramatic Flower Color Polymorphism in the Rewardless Orchid *Dactylorhiza sambucina*", *Proceedings of the National Academy of Sciences, USA* 98 (6253-6255).

Gildenhuys, P. (forthcoming), "Causal Equations without *Ceteris Paribus* Clauses", *Philosophy of Science*.

Gildenhuys, Peter (2003), "Darwin, Herschel, and the Role of Analogy in Darwin's Origin", *Studies in History and Philosophy of Biological and Biomedical Sciences* 35:593-611.

——— (2008), "An Explication of the Causal Dimension of Drift", *British Journal for the Philosophy of Science* [forthcoming] http://tinyurl.com/4l4e7f.

Gillespie, John (1991), *The Causes of Molecular Evolution*. Oxford: Oxford University Press.

——— (1998), *Population Genetics: A Concise Guide*. Baltimore, Maryland: The Johns Hopkins University Press.

Glymour, Bruce (2006), "Wayward Modeling: Population Genetics and Natural Selection", *Philosophy of Science* 73:369-389.

Glymour, Clark, Richard Scheines, and Peter Spirtes (1993), *Causation, Prediction, and Search*. Cambridge, MA: MIT Press.

Godfrey-Smith, Peter (2000), "The Replicator in Retrospect", *Biology and Philosophy* 15 (3):403-423.

Godfrey Smith, Peter (2000), "The Replicator in Retrospect", *Biology and Philosophy* 15 (3):403-423.

Godfrey Smith, Peter, and Benjamin Kerr (2002), "Individualist and Multi-level Perspectives on Selection in Structured Populations", *Biology and Philosophy* 17:477-517.

Gould, Stephen J. (2002), *The Structure of Evolutionary Theory*. Cambridge: Harvard University Press.

Gould, Stephen J., and R. C. Lewontin (1979), "The Spandrels of San Marco and the Panglossian Paradigm", *Proceedings of the Royal Society of London* 205 (1161):581-598.

Griesemer, James (2000), "Development, Culture, and the Units of Inheritance", *Philosophy of Science* 67 (Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers):S348-S368.

——— (2000), "The Units of Evolutionary Transition", *Selection* 1:67-80.

Griffiths, P. E., and R. D. Gray (1994), "Developmental Systems and Evolutionary Explanation", *The Journal of Philosophy* 91 (6):277-304.

——— (2004), "The Developmental Systems Perspective: Organism-environment systems are units of evolution", in K Preston and M Pigliucci (eds.), *The Evolutionary Biology of Complex Phenotypes*, New York: Oxford University Press, 1-23.

Griffiths, P. E., R. D. Gray, and Susan Oyama (2001), "What is Developmental Systems Theory?" in Susan Oyama, P. E. Griffiths and R. D. Gray (eds.), *Cycles of Contingency*, Cambridge, MA: MIT Press.

Griffiths, P. E., and Russell D. Gray (2001), "Darwinism and Developmental Systems", in Susan Oyama (ed.), *Cycles of Contingency*, Cambridge, MA: MIT Press, 195-218.

Griffiths, P. E., and Eva M. Neumann-Held (1999), "The Many Faces of the Gene", *Bioscience* 49 (8):656-662.

Griffiths, Paul E., and Russell D. Gray (1997), "Replicator II--Judgement Day", *Biology and Philosophy* 12 (4):471-492.

Griffiths, Paul E., and Robin D. Knight (1998), "What Is the Developmentalist Challenge?" *Philosophy of Science* 65 (2):253-258.

Hamilton, W.D. (1964), "The Genetical Evolution of Social Behavior I and II", *Journal of Theoretical Biology* 7:1-16; 17-52.

Hanski, I, and M Gilpin, eds. (1996), *Metapopulation Dynamics, Ecology, Genetics, and Evolution*. New York: Academic Press.

Hartl, Daniel, and Andrew Clark (2007), *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates.

Hartl, Daniel., and Andrew Clark (1988), *Principles of Population Genetics*. Sunderland, MA: Sinauer Associated.

Hedrick, Philip W. (1990), "Theoretical Analysis of Habitat Selection and the Maintenance of Genetic Variation", in J.S.F. Barker, W.T. Starmer and R. McIntyre (eds.), *Ecological and Evolutionary Genetics of Drosophila*, New York: Plenum.

——— (1993), "Sex-Dependent Habitat Selection and Genetic Polymorphism", *American Naturalist* 141 (3):491-500.

——— (2000), *Genetics of Populations*. Second ed. Sudbury, MA: Jones and Bartlett Publishers.

——— (2005), *Genetics of Populations*. Third ed. Boston: Jones and Bartlett.

Heinrich, Joseph, R. Boyd, and P.J. Richerson (2008), "Five Misunderstandings about Cultural Evolution", *Human Nature* 19:119-137.

Heinrich, Joseph, and Robert Boyd (2002), "On Modeling Cognition and Culture: Why cultural evolution does not require replication of representations", *Journal of Cognition and Culture* 2 (2):87-112.

Hori, Michio (1993), "Frequency-Dependent Natural Selection in the Handedness of Scale-Eating Cichlid Fish", *Science* 260:216-219.

Huffaker, C.B., A.A. Berryman, and J.A. Laing (1984), *Ecological Entomology*. New York: Academic Press.

Hull, David L. (1980), "Individuality and Selection", *Annual Review of Ecology and Systematics* 11:311-332.

——— (1988), *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

——— (2001), *Science and Selection: Essays on biological evolution and the philosophy of science*. Cambridge: Cambridge University Press.

Hull, David L., Sigrid Glenn, and Rod Langman (2001), "A General Analysis of Selection", in David L. Hull (ed.), *Science and Selection*, Cambridge: Cambridge University Press.

Human, Kathleen, and Deborah Gordon (1996), "Exploitation and Interference Competition between the Invasive Argentine Ant, *Linepithema humile*, and Native Ant Species", *Oecologia* 105:405-412.

Jacquard, Albert (1974), *The Genetic Structure of Populations*. New York: Springer-Verlag.

Keller, Laurent, and Kenneth Ross (1998), "Selfish genes: a green beard in the red fire ant", *Nature* 395:572-575.

Kerr, Benjamin, and Peter Godfrey-Smith (2002), "Individualist and Multi-level Perspectives on Selection in Structured Populations", *Biology and Philosophy* 17:477-517.

Kerr, Benjamin, and Peter Godfrey Smith (2008), "Generalization of the Price Equation for Evolutionary Change", *Evolution* 63 (2):531-536.

Kimura, M. (1965), "Attainment of Quasi Linkage Equilibrium when Gene Frequencies Are Changing by Natural Selection", *Genetics* 52 (5):875-890.

Kirkpatrick, Mark, Toby Johnson, and N. H. Barton (2002), "General Models of Multilocus Evolution", *Genetics* 161:1727-1750.

Kitcher, Philip (1985), "Darwin's Achievement", in Nicholas Rescher (ed.), *Reason and Rationality in Natural Science*, Boston, MA: University Press of America, Inc.

Kitcher, Philip, and K. I. M. Sterelny (1988), "The Return of The Gene", *Journal of Philosophy* 85:339-361.

Lanning, Dirk V., and James T. Shiflett (1983), "Nesting Ecology of Thick-Billed Parrots", *The Condor* 85 (1):66-73.

Lapedes, D.N., ed. (1978), *McGraw-Hill Dictionary of Scientific and Technical Terms*. 2nd ed. New York: McGraw-Hill.

Leigh, Egbert Giles (1999), "Levels of Selection, Potential Conflicts, and Their Resolution: The Role of the "Common Good" ", in H.K. Reeve and E. Keller (eds.), *Levels of Selection in Evolution*, Princeton, NJ: Princeton University Press, 15-30.

Lennox, James (1991), "Darwinian Thought Experiments: A Function For Just So Stories", in T. Horowitz and G. Massey (eds.), *Thought Experiments in Science and Philosophy*, Savage, Md.: Rowman and Littlefield, 173-195.

Lennox, James, and Bradley E. Wilson (1994), "Natural Selection and the Struggle for Existence", *Studies in History and Philosophy of Science* 25 (1):65-80.

Levene, H. L. (1953), "Genetic Equilibrium When More Than One Ecological Niche is Available", *The American Naturalist* 87 (836):331-333.

Levins, Richard (2004), "Toward a Population Biology, Still", in Rama S. Singh and Marcy K. Uyenoyama (eds.), *The Evolution of Population Biology*, Cambridge: Cambridge University Press, 21-47.

Lewontin, R. C. (1970), "The Units of Selection", *Annual Review of Ecology and Systematics* (1):1-18.

——— (1974), *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

——— (1978), "Adaptation", *Scientific American* 239 (3):213-230.

——— (1991), *Biology as Ideology*. New York: Harper.

Lloyd, Elisabeth (1992), "Unit of Selection", in Elisabeth Lloyd and E. Keller (eds.), *Keywords in Evolutionary Biology*, Cambridge: Harvard University Press.

Lyell, C. (1969[1830]), *Principles of Geology*. New York: Johnston Reprint Company.

Matthen, Mohan, and Andre Ariew (2002), "Two Ways of Thinking about Fitness and Natural Selection", *The Journal of Philosophy* XCIX (2):55-83.

——— (2005), "How to Understand Causal Relations in Natural Selection: Reply to Rosenberg and Bouchard", *Biology and Philosophy* 20:355-364.

Maynard-Smith, John (1983), "Models of Evolution", *Proceedings of the Royal Society of London B* 219:315-325.

Maynard Smith, John (1964), "Group Selection and Kin Selection", *Nature* 201:1145-1147.

——— (1983), "Models of Evolution", *Proceedings of the Royal Society of London B* 219:315-325.

——— (1987), "How to Model Evolution", in John Dupre (ed.), *The Latest on the Best*, Cambridge, Mass,: MIT Press.

——— (1988), "Evolutionary Progress and Levels of Selection", in Matthew H. Nitecki (ed.), *Evolutionary Progress*, Chicago: University of Chicago Press, 219-230.

——— (1991), "A Darwinian View of Symbiosis", in Lynn Margulis and René Fester (eds.), *Symbiosis as a Source of Evolutionary Innovation*, Cambridge, MA: MIT Press, 26-39.

Maynard Smith, John, and Szathmáry (1995), *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Mendel, Gregor (1901[1865]), "Experiments in Plant Hybridization".

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

Millstein, Roberta (1996), "Random Drift and the Omniscient Viewpoint", *Philosophy of Science* 63:S10-S18.

——— (2002), "Are Random Drift and Natural Selection Conceptually Distinct?" *Biology and Philosophy* 17 (1):33-53.

——— (2005), "Selection vs. Drift: a response to Brandon's reply", *Biology and Philosophy* 20 (1):171-175.

Moxon, Richard, Paul Rainey, Martin Nowak, and Richard Lenski (1994), "Adaptive Evolution of Highly Mutable Loci in Pathogenic Bacteria", *Current Biology* 4 (1):24-33.

Nanay, Bence (2002), "The Return of the Replicator: What is Philosophically Significant in a General Account of Replication and Selection", *Biology and Philosophy* 17:109-121.

Okasha, Samir (2006), *Evolution and the Levels of Selection*. New York: Oxford University Press.

Pearl, Judea (2000), *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Plutynski, Anya (2004), "Explanation in Classical Population Genetics", *Philosophy of Science* 71:1201-1214.

Rice, Sean (2004), *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, MA: Sinauer and Associates.

Robertson, A. (1962), "Selection for Heterozygotes in Small Populations", *Genetics* 47:1291-1300.

Rosenberg, Alexander, and F. Bouchard (2004), "Fitness, Probability and the Principles of Natural Selection", *British Journal for the Philosophy of Science* 55:693-712.

Rosenberg, Alexander, and Frederic Bouchard (2005), "Matthen and Ariew's Obituary for Fitness: Reports of its Death Have Been Greatly Exaggerated", *Biology and Philosophy* 20:343-353.

Rosenzweig, Michael L. (2003), *Win-win Ecology: How the Earth's Species can Survive in the Midst of Human Enterprise*. New York: Oxford University Press.

Roughgarden, J. (1971), "Density-Dependent Natural Selection", *Ecology* 52 (3):453-468.

Salmon, Wesley C. (1998), *Causality and Explanation*. Oxford: Oxford University Press.

Santiago, E., and Armando Caballero (1998), "Effective Size and Polymorphism of Linked Neutral Loci in Populations under Directional Selection", *Genetics* 149:2105-2117.

Schemske, Douglas W., and Paulette Bierzychudek (2001), "Perspective: Evolution of Flower Color in the Desert Annual Linanthus parryae: Wright revisited", *Evolution* 55 (7):1269-1282.

——— (2007), "Spatial Differentiation for Flower Color in the Desert Annual *Linanthus Parryae*: Was Wright Right?" *Evolution* 61 (11):2528-2543.

Seger, J (1985), "Unifying Genetic Models for the Evolution of Female Choice", *Evolution* 39:1185-1193.

Shanahan, Timothy (2003), "The Evolutionary Indeterminacy Thesis", *Bioscience* 53 (2):163-169.

Smith, T. B. (1993), "Disruptive Selection and the Genetic Basis of Bill Size Polymorphism in the African Finch *Pyrenestes*", *Nature* 363:619-620.

Sober, Elliott (1984), *The Nature Of Selection: Evolutionary Theory In Philosophical Focus*. Cambridge, Mass.: MIT Press.

——— (1987), in John Dupre (ed.), *The Latest on the best : essays on evolution and optimality* Cambridge, Mass.: MIT Press, 359.

Sober, Elliott, and Richard C. Lewontin (1982), "Artifact, Cause, and Genic Selection", *Philosophy of Science* 49 (2):157-180.

Sober, Elliott, and David Sloan Wilson (1994), "A Critical Review of Philosophical Work on the Units of Selection Problem", *Philosophy of Science* 61 (4):534-555.

——— (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge: Harvard University Press.

Stahl, Eli A., Greg Dwyer, Rodney Mauricio, Martin Kreitman, and Joy Bergelson (1999), "Dynamics of Resistance Polymorphism at the *Rpm1* locus of *Arabidopsis*", *Nature* 400:667-671.

Sterelny, Kim, Kelly C. Smith, and Michael Dickison (1996), "The Extended Replicator", *Biology and Philosophy* 11 (3):377-403.

Suppe, Frederick (1974), "Introduction", in Frederick Suppe (ed.), *The Structure of Scientific Theories*, Urbana: University of Illinois Press, 3-243.

Szathmáry, Eors (1999), "The First Replicators", in Laurent Keller (ed.), *Levels of Selection in Evolution*, Princeton: Princeton University Press.

Templeton, A.R., and E.D. Rothman (1981), "Evolution in Fine-Grained Environments", *Evolution* 38:596-608.

Turelli, Michael, Douglas W. Schemske, and Paulette Bierzychudek (2001), "Stable Two-Allele Polymorphisms Maintained by Fluctuating Fitnesses and Seed Banks: Protecting the Blues in Linanthus parryae", *Evolution* 55 (7):1283-1298.

Wallace, Bruce (1968), "Polymorphism, Population Size, and Genetic Load", in R. C. Lewontin (ed.), *Population Biology and Evolution*, Syracuse: Syracuse University Press, 87-108.

——— (1975), "Hard and Soft Selection Revisited", *Evolution* 29 (3):465-473.

Walsh, D. M. (2004), "Bookkeeping or Metaphysics? The Units of Selection Debate", *Synthese* 138 (3):337-361.

——— (2007), "The Pomp of Superfluous Causes", *Philosophy of Science* 74:281-303.

Walsh, Denis M., Tim Lewens, and Andre Ariew (2002), "The Trials of Life: Natural Selection and Random Drift", *Philosophy of Science* 69 (3):452-473.

Waples, Robin S., and Oscar Gaggiotti (2006), "What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity", *Molecular Ecology* 15:1419-1439.

Wedekind, C., T. Seebeck, F. Bettens, and A Paepke (1995), "MHC-dependent mate preferences in human", *Proceedings of the royal Society of London B.* 260:245-249.

Williams, George C (1966), *Adaptation and Natural Selection*. Princeton, N.J.: Princeton University Press.

Wilson, David Sloan (1990), "Weak Altruism, Strong Group Selection", *Oikos* 59 (1):135-140.

Wimsatt, William C. (1980), "The Units of Selection and the Structure of the Multi-Level Genome", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1980 (2):122-183.

Woodward, James (2003), *Making Things Happen*. New York: Oxford University Press.