

Neutral Theory, Biased World

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

William C. Bausman

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: C. Kenneth Waters

December 2016

Acknowledgements

First, I want to thank two people without whose support this dissertation would never have been finished:

Ken Waters, who gave me the initial push to extend a great seminar discussion into a seminar paper, and then into my dissertation. Ken's support has never faltered, and I have felt better and more focused after every meeting we have ever had.

Past-Will, who faltered many times, but who did all the work here. I could not be writing this without him.

Next, I want to thank the members of my dissertation committee:

Steve Hubbell, for giving me such a great case study to write my dissertation on, and who always had time for a conversation with me.

Alan Love, for always saying back to me what I was trying to say much more clearly than I could say it myself.

Bill Wimsatt and Jos Uffink, for being exemplary scholars and role models.

Finally, I want to thank:

The members of BIG, where I learned how to do philosophy of science.

Marta Halina, for all the hours talking about and writing what became chapter 3.

My friends Melanie Bowman, Shay Logan, Andrew Inkpen, and Jack Powers, for our conversations.

My parents, who always encouraged me in everything.

My partner Katharina Steiner, for her friendship, love, and support.

Table of Contents

Acknowledgements.....	i
Table of Contents.....	ii
Table of Figures.....	iii
Table of Tables.....	iv
Introduction.....	1
Chapter 1.	6
Telling the Origins of the Neutral Theory of Ecology	
Chapter 2.	69
Retelling the Origins	
Chapter 3.	130
Not Null Enough: Non-Statistical Null Hypotheses in Community Ecology and Comparative Psychology	
Chapter 4.	153
Modeling: Neutral, Null, and Baseline	
Chapter 5.	183
False Models for Fecund Research Programmes	
References.....	237

List of Figures

Figure	Page
Figure 1.1 – Death, Birth, Immigration cycle in Neutral Theory.....	11
Figure 1.2 – Species-Area plot of Sunda Islands.....	20
Figure 1.3 – Species-Area curve of West Indian Herpetofauna.....	21
Figure 1.4 – Island Biogeography Model.....	22
Figure 1.5 – Variation in Island Biogeography models.....	23
Figure 1.6 – Lognormal distribution in different bases.....	34
Figure 1.7 – Comparison of species abundance distributions.....	35
Figure 1.8 – Lognormal Distribution fit to data with veil line.....	40
Figure 1.9 – Lognormal Distribution fit to data with veil line.....	40
Figure 1.10 – Comparison of Whittaker and Preston plots.....	41
Figure 1.11 – Comparison of Geometric, Log-series, and Log-normal Distributions.....	42
Figure 1.12– Relationship between Geometric, Log-series, and Log-normal Distributions.....	43
Figure 1.13 – Broken-Stick Model on Whittaker Plot.....	47
Figure 1.14 – Comparison of Niche Apportionment Model Distributions....	51
Figure 2.1 – Incidence Functions of Cuckoo-doves on Bismarck Archipelago.....	111
Figure 4.1 – Comparison of species abundance distributions.....	169
Figure 4.2 – Death, Birth, Immigration cycle in Neutral Theory.....	171
Figure 4.3 – Close fit of Neutral Model to BCI SAD.....	173
Figure 5.1 – Comparison of species abundance distributions.....	221

List of Tables

Table	Page
Table 2.1 – Relationship between three sets of distinctions: Strong vs Weak, Null Hypothesis Testing vs. Null modeling vs. baseline Modeling, and Statistical Null Hypothesis vs. Dynamical Null Hypothesis	124

Introduction

This dissertation investigates how community ecology can inform history and philosophy of science, and how history and philosophy of science can inform community ecology. My case study is Stephen Hubbell's neutral theory of ecology. The neutral theory is a prominent and controversial theory, and the ongoing debates surrounding it raise deep philosophical questions. Nevertheless, it has received little philosophical attention.

This work is organized so that each chapter can be read independently while they progressively build to chapter 5. I aim to explain the philosophical issues raised by the practices of ecologists when they develop and use the neutral theory. I take a historical approach to characterizing and informing these issues by putting them in the context of both the history of biology since the 1970s, and the philosophy of science since Karl Popper. I use the philosophical methods of analyzing both the explicit and implicit argumentation used, and drawing relevant distinctions to lay bare epistemic commitments and expose the reasoning used so that it can be evaluated and improved.

Chapter 1. Telling the Origins of the Neutral Theory of Ecology

Where did the neutral theory come from? I analyze Stephen Hubbell's answer to this question as a scientific *origin story*. I first explain and critique Hubbell's origin story historically and show how the neutral theory developed as a formal model of patterns of abundance and diversity. I argue that the strength of this origin story is as a conceptual presentation of the theory, but that its weakness is that it represents as a eureka moment what actually took place over many years. I then expand Hubbell's origin story into a

broader history detailing the punctuated development of formal models in community ecology. In addition to being a historical inquiry, I use this chapter to frame the received view of the epistemological issues raised by the neutral theory.

Chapter 2. Retelling the Origins

This chapter investigates two additional historical questions about the origins of the neutral theory. First, What was Hubbell working on before the neutral theory? Answering this question reveals that the original context of the neutral theory is in tropical forest ecology, and that the original problem addressed is the taxonomic equilibrium problem – is there species turnover in an otherwise stable community? Hubbell argued, using the ancestor of the neutral theory, that tropical forests are not in equilibrium. This narrative serves, among other things, to show that the neutral theory was not originally an abstract and general theory aimed at explaining patterns of abundance and diversity.

Second, Why did Hubbell call his theory “neutral”? Hubbell only adopted the lasting title of “neutral theory” shortly before publishing his 2001 monograph. I show that he did so because he was pushed by others to acknowledge that, as a *neutral* theory, it was linked to particular methods for investigating complex systems, variously called *null hypothesis testing* and *null modeling*. These methods have a lively and controversial history in ecology, within which I situate the most interesting epistemic issues raised by the neutral theory. These two answers together serve to re-frame the epistemological questions raised by the practice of ecologists using the neutral theory. I take these up in the next three chapters.

Chapter 3. Not Null Enough: Non-Statistical Null Hypotheses in Community

Ecology and Comparative Psychology

This chapter is co-authored with Marta Halina, now lecturer at Cambridge HPS. We contributed equally to this work. This chapter identifies and evaluates a common reasoning strategy used in community ecology and comparative psychology for selecting between competing hypotheses. This strategy labels one hypothesis as a “null” on the grounds of its causal simplicity and proceeds to grant it epistemic privileges on the grounds that it is a statistical null hypothesis. We argue that this strategy is unjustified in the cases of the neutral theory of ecology and the behavioral-rules hypothesis in comparative psychology because these cannot properly function as statistical nulls. The use of statistical nulls is justified through the experimental and mathematical contexts in which they are used, but these contexts are missing in our two case studies. This reasoning strategy should be understood as an unjustified rhetorical strategy used to gloss over a commitment to valuing simplicity more than other epistemic virtues in the name of good scientific, statistical methodology.

Chapter 4. Modeling: Neutral, Null, and Baseline

This chapter takes up where chapter 3 leaves off. Both chapters critique the use of the neutral theory as a *null* with respect to competition theory. Determining whether this use of the neutral theory establishes its conclusions depends upon how one understands the reasoning strategy “null” refers to. The following encapsulates this reasoning in question:

1. The neutral theory supplies the appropriate null with respect to competition theory.
2. The neutral theory is empirically adequate for describing some patterns of interest.
3. Therefore, the neutral theory accurately identifies the dominant causal processes responsible for the pattern of interest.

Chapter 3 analyzes this line of argumentation's use of "null" as referring to the reasoning strategy of *null hypothesis testing*, in which a statistical null hypothesis is used to test for whether there is a pattern to be explained. There we argue that, because the neutral theory does not supply a statistical null hypothesis, it should not be analyzed as such.

But there are two additional reasoning strategies that might be intended by "null": *null modeling* and *baseline modeling*. Null modeling uses a model as null to test for whether some proposed process is causally relevant to a pattern. Baseline modeling uses a model as null to apportion relative significance to multiple processes, each of which is assumed to be relevant to a pattern. I show that the neutral theory does supply appropriate models for both null modeling and baseline modeling. However, the line of argumentation schematized above still does not follow using either reasoning strategy. If null modeling is intended, then the reasoning is invalid because null modeling can only grant evidence to the model being tested by the null model. If baseline modeling is intended, then the reasoning is valid but unsound because there is no justification given for the neutral theory supplying the appropriate baseline model.

The most charitable reading of the neutralists reasoning is that they intend to be using baseline modeling. However, supplying the appropriate baseline model requires an independent and substantial justification that cannot come from null modeling and merely

assuming neutrality, and which has yet to be made clear by the neutralists.

Chapter 5. False Models for Fecund Research Programmes

Scientists sometimes enlist philosophy of science when they are engaged in scientific controversies. Many ecologists endorse Karl Popper's philosophy of science. For example, Stephen Hubbell has claimed that the neutralists' approach is more scientific than its rivals because the neutralists use the neutral theory qua null to *test hypotheses*, while the rivals are doing merely *confirmatory science*.

But the Popperian image of science as conjecture and refutation of theories and hypotheses is impoverished because cannot capture the methodological dimensions of the scientists working with the neutral theory. I show that the epistemology of the neutral theory is better understood using Lakatos' concept of a *research programme*. I show that broadening philosophical attention to the *neutralist research programme* illuminates (1) the synchronic uses of the neutral theory to make predictions and give descriptions and explanations; (2) its diachronic development in response to theoretical innovation and confrontation with data; (3) its complex relationships to theories based on competition across differentially adapted species. For example, the reasoning strategy *baseline modeling* turns out to be the primary synchronic heuristic in the neutralist research programme. The justification for using the neutral theory in baseline modeling, found wanting in chapter 4, should be sought instead in the neutralist research programme. This chapter completes the re-framing of the epistemology set up in chapters 1 and 2.

Chapter 1.

Telling the origins of the neutral theory of ecology

1. The Neutral Theory of Ecology

Stephen Hubbell published *The Unified Neutral Theory of Biodiversity and Biodiversity* in 2001 (Hubbell 2001). In it, Hubbell presented the basic version of what I will refer to as *the neutral theory of ecology*, or *neutral theory* for short. The basic neutral theory has both antecedent and subsequent developmental stages, many of which will be discussed in this dissertation. But the basic version is the best version for understanding what the theory hypothesizes, how the basic model works, and why it is controversial.

The neutral theory of ecology is a theory of biodiversity and biogeography in ecological communities. More specifically, it is a theory of *single trophic level* communities. A single trophic level community contains only organisms that eat, and are eaten by, the same kinds of things. The first trophic level are the primary producers, the plants that get their energy from the sun. The second trophic level are the primary consumers, those organisms that eat the plants. And so on. For example, all of the trees in a forest are studied as a single trophic level community. All the birds on an island, excluding the raptors, are also studied as a single trophic level community. Examples of single trophic level communities are all the trees in a forest, and all the birds on an island excluding the raptors. The empirical targets of the neutral theory are patterns of communities including *relative species abundance distributions* (SAD), which record the number of species and number of individuals in each species in a community. Other

patterns targeted by the neutral theory include: the *species-area relationship* (SAR), which records how the number of species in a community changes with the physical area of the community; and patterns describing how SADs and SARs change with spatial and temporal scale and vary across communities.

The neutral theory aims to describe, predict, and explain these empirical patterns. To these ends, the neutral theory hypothesizes that these patterns are the result of three processes: ecological drift, immigration, and speciation. These demographic and dispersal processes control the flow of individuals in and out of a community. The neutral theory makes two formal assumptions. First, the neutral theory assumes *saturation*, that the size of the community is fixed over time at a finite number of individuals. Assuming saturation is meant to represent that resources are fixed and that none go unused for long in mature and undisturbed communities. Second, the neutral theory assumes *neutrality*, that all the individuals in the (single trophic level) community are functionally equivalent in the sense that there are no relevant differences between individuals of different species. For example, in a neutral community with only birth and death, each individual is equally likely to die, and each remaining individual is equally likely to successfully reproduce. Together then, the neutral theory hypothesizes that groups of organisms compete as equals for space through random drift, random immigration, and random speciation. When one individual dies, another takes its place – either from reproduction, immigration, or speciation. Because of neutrality, the probability that a given species will fill a gap is just proportional to the present abundances.

The neutral theory is formalized using mathematics to construct the basic neutral model, from which predictions can be derived. The basic neutral model has 3 free,

biologically-interpreted parameters: J , community size; m , immigration rate; and θ , called the fundamental biodiversity number. The parameter θ is defined as a function of two other parameters: J_M , metacommunity size; and ν , speciation rate. The neutral theory uses its hypotheses about the target patterns and the neutral model to predict that the target patterns will be statistically indistinguishable from specific statistical distributions.

To get an intuitive understanding of the neutral theory, it is helpful to have an informal picture of how the neutral model functions. Because the neutral model, itself a mathematical object, is being considered as a model of patterns of abundance in communities, the model will be described under the ecological interpretation given to it by the neutral theory. This means that the same terms (such as local community, drift, etc.) will be used to describe both the model and the system modeled. The primary process in the neutral theory is *ecological drift*, also called demographic stochasticity or random birth and death. To understand how a purely drifting community behaves, imagine an island community as a checkerboard populated by differently colored pieces.¹ We will track how the relative abundance of colored pieces change over time by demographic and dispersal processes. I use this checkerboard depiction to show the general behavior of the neutral model and why immigration and speciation are also included. I also use it to draw out the general remarks about the neutral theory needed to

¹ Picturing the neutral model of ecology as a checkerboard has been used in (Purves and Pacala 2005) and (Etienne and Haegeman 2011). Different sets of rules for the checkerboard game can be given to explain different models. This has been done extensively in evolution for both selection and neutral models. Beatty (Beatty 1984) discusses these in the context of Kimura's neutral theory of molecular evolution, and locates one important source of the checkerboard framework in (Eigen and Winkler 1981).

understand the historical and philosophical issues examined in what follows.

To begin our informal depiction, imagine that you have a checkerboard and a bag of pieces of various colors. Fill the board with pieces drawn blindly from the bag. The board corresponds to the island or *local community*, and the bag correspond to the mainland or *metacommunity*. Record the relative number of pieces of each color on the board. For example, 5 reds, 4 greens, 2 yellows, and 1 each of blue, purple, and orange. You might depict this using a pie chart. This data makes up the initial SAD of the board. Our goal is to determine the time-independent equilibrium SADs by explaining the possible dynamics of SADs. This is done by introducing rules for deaths and replacements. In order to satisfy saturation, the size of the board must contain a finite number of spaces, and any space opened by a death must be immediately filled. To begin, remove a piece according to the *death rule*: each piece currently on the board has an equal chance of being removed.

Births are the primary way to balance out deaths and maintain community size. Imagine that the pieces on the board can reproduce. Under neutrality, each piece on the board has an equal chance of reproducing. Therefore, the probability of successful birth by color p is proportional to the relative abundance of p on the board. This is the *drift rule*. Extinction is not an independent process in the neutral theory, but is a product of drift. With only drift, over time the number of species can never increase and will tend to decrease, eventually leading to only one species occupying the whole community. Hence, the only time-independent equilibrium state under pure drift is where one species occupies the whole community.

To construct an equilibrium model of *biodiversity*, the possibility of an individual immigrating to our community from the bag (metacommunity) is incorporated. Instead of replacing every death with a reproduction according to the drift rule, now there is an *immigration probability* m that an individual will immigrate from the bag. Following every death and opening on the board, with probability $(1-m)$, follow the drift rule and fill the gap with a piece of a given color with probability proportional to current relative abundance of colors on the board. And with probability m , replace the removal with an immigrant from the bag of pieces following the immigration rule. Analogous to the drift rule, the *immigration rule* says that the probability of immigration by color q is proportional to the relative abundance of q in the bag. In this way, the SAD of a local community is determined by the balance of drift (including random deaths) and immigration.

Altogether then, the SAD of the board evolves over time through the following steps:

1. Remove a piece following the death rule, and go to step 2.
2. Fill the gap on the board with either step 2.1 or 2.2, and then go to step 1,
 - 2.1. With probability $1-m$, by a piece on the board reproducing following the drift rule,
 - 2.2. With probability m , with a piece from the bag immigrating following the immigration rule.

This iterative process of death and either local birth or immigration are represented graphically in figure 1.1.

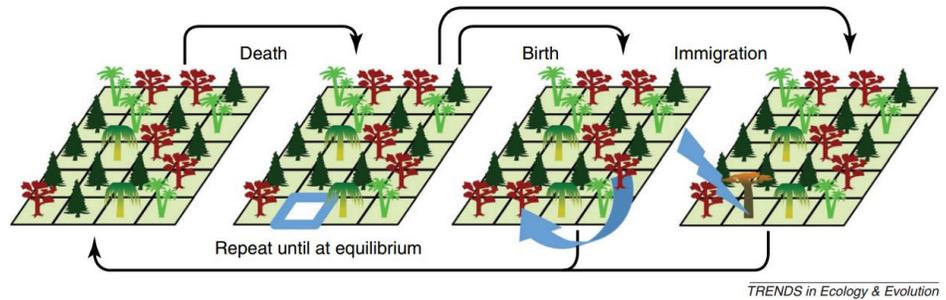


Figure 1.1 – Death, Birth, Immigration Cycle in Neutral Theory
Figure from (Rosindell, Hubbell, and Etienne 2011). Used with permission from Elsevier.

If individuals are immigrating to the local community from the metacommunity, and the probability of a particular individual immigrating is determined by the relative abundance of that species in the metacommunity, then the relative abundances in the metacommunity also needs to be tracked. Therefore, assume that the bag is also fixed at a finite size and that drift is also occurring in the bag. Then pieces are dying and being replaced by reproduction according to the same drift rule. With finite size and only drift and emigration, eventually the bag will lose all of its color diversity, and this will in turn lead the local community to lose its color diversity.

In order to construct a theory that can maintain a stable equilibrium of any true diversity then, introduce a process of speciation in the metacommunity to serve the same function that immigration does for the local community. Speciation is the ultimate source of diversity in the theory, with immigration carrying the diversity to the local communities and drift slowly tending each species go extinction. Speciation can be modeled in a number of ways and so there are many different neutral models, each with a different speciation mechanism. The simplest version of speciation works like a point mutation where a new species pops up in an instant, imagined as a piece with a novel color appearing in the bag, in response to a death.

When the rounds of death and replacement are repeated enough times, the time-independent equilibrium SAD will be produced. Remember that the SAD of a community records the number of individuals in each species. The time-independent SAD depends only upon the values of the free parameters, not the initial SAD whose trace is washed away over time. First, the equilibrium SAD for the bag/ metacommunity is determined by the value of the fundamental biodiversity number θ , which is defined as a function of two other parameters: metacommunity size J_M and speciation rate v . Next the equilibrium SAD for the board/ local community is determined by the values of θ and also local community size J and immigration probability m .

The time-independent equilibrium solutions of the neutral model which are, for the most part, compared with SAD data from communities. The dynamics of the near-equilibrium and far-from-equilibrium behavior of the model are not well understood. I have only described the model with respect to SADs, but the same game can be used to understand how the neutral model makes predictions about the other patterns such as the species-area relationship (SAR), which record how the number of species scales with the size of the community. SARs are predicted by running the model in different sizes and recording the equilibrium number of species for each community size.

Notice that, by assuming neutrality, the board game analogy explicitly excludes the possibility that pieces of different colors have different chances of reproducing, immigrating, and speciating. These kinds of differences would lead to competition between colors (*interspecific competition/ selection*) for the resource of space, and the colors with higher chances would tend to increase their relative abundances at the expense of the less well-adapted colors. Introducing selective differences between species

would lead to equilibrium SADs that deviate more or less from the expectations of the neutral model sketched here. The size of the deviation depends on the size of the board, the bag, the immigration probability, the speciation probability, and the selective differences. This is because, by assuming neutrality, the neutral theory hypothesizes that interspecific competition/ selection is not relevant to patterns of abundance such as SADs and SARs in real communities.

With the checkerboard understanding of the neutral model and the general understanding of the neutral theory gleaned from it, we can now begin the historical and philosophical investigation of the neutral theory.²

2. A Scientific Controversy

Stephen Hubbell ignited a controversy within community ecology when he published *The Unified Neutral Theory of Biodiversity and Biogeography* in 2001. One reviewer of the book identified the source of the controversy in the following way,

“This neutral theory has already sparked some controversy in the literature and has inspired many studies. Although the theory’s predictions seem consistent with much empirical data, how can such a theory, with assumptions that are so obviously wrong, be useful?” (de Mazancourt 2001)

The ‘obviously wrong’ assumption is *neutrality*, that all individuals in a single tropic level community, regardless of species, are ecologically equivalent. Even though the neutral theory is capable of predictions consistent with empirical observations about the patterns of abundance and diversity, neutrality is seen as obviously false because it

² Readers seeking for mathematical detail should first consult the following: (Hubbell 2001), (Volkov et al. 2003), (Etienne and Alonso 2007).

conflicts with the explanation of these patterns with interspecific competition. The dominant way ecologists explain why a community of nesting birds has such and such many members of each species is to cite the causally relevant *differences* between how different species utilize resources and the relative abundances of the particular resources in that community. The patterns of abundance and diversity were hypothesized to be the result of competition across different species for resources. But by assuming neutrality, the neutral theory hypothesizes that there are no such relevant differences between species, and that instead the relevant causal factors are the *similarities*. Stephen Hubbell himself has expressed the same view of the source of the controversy,

“The hypothesis of ecological equivalence is the fundamental yet controversial idea behind neutral theory.” (Hubbell 2006, 1387)

The philosopher of science immediately recognizes this as a general epistemological problem (see, for example, Odenbaugh’s framing in (Odenbaugh Forthcoming)). We might express it as the following tension:

- A. The empirical success of a theory implies the theory is useful/ explanatory.
- B. Making an obviously false assumption implies the theory is not useful/ explanatory.

Expressed in this way, the controversy surrounding the neutral theory has the following characteristics:

1. It is a particular case of the general problem of understanding how false assumptions can be useful to science.
2. It takes the scientists to be working on the problem of inferring process from pattern based on the goodness of fit between models and data.

3. It portrays the controversy as between competition and neutral worldviews.
4. It gives the impression that the neutral theory takes a very abstract approach to modeling complex systems, informed more by mathematics than ecology.

In this way, attention to the neutral theory has focused on the model-fit relationship and the inferences licensed from this, especially given that it assumes neutrality, which conflicts with the dominant explanation. This has proven to be an attractive characterization of the situation to ecologists who are fixated on testing hypotheses both alone and in comparison to alternatives. It is also attractive to epistemologists and philosophers of science who have a predilection for analyzing scientific controversies in terms of theories, opposed worldviews, and in terms of general problems. But such an understanding of the use and the controversy surrounding the neutral theory's use is too course-grained. It limits our understanding of the theory, its targeted patterns, the reasoning strategies used, and of ecology, biology, and the sciences generally. How are we to change our understanding of the issues raised by the neutral theory?

Histories are told to inform and support the way we view things presently and want them to change in the future. The way the neutral theory is framed and addressed is dependent upon the historical narrative told about its origins. Stephen Hubbell tells a particular story about the origins and development of the neutral theory of ecology which foregrounds the four-fold characterization given above, but backgrounds others. In this way, the existing philosophy of the neutral theory, told both by ecologists themselves and philosophers of ecology, has been founded upon its history. But that history is not the

only one that can be told. And we can change the way the epistemology is approached by changing the way we tell the origins of the neutral theory.

The received history, based on the way that Hubbell tells the neutral theory's origins, informs the current discussion of the neutral theory both among ecologists and philosophers. This chapter and the next are devoted to presenting the received history of the neutral theory and then answering three different historical questions about the origins of the neutral theory. I begin in section 3 by presenting the story which Hubbell tells of the origins of the neutral theory in his 2001 book. Section 4 addresses Hubbell's origin story as historical claim and motivates digging deeper into the origins of the neutral theory. The first narrative, presented in section 5, is the received history of the theory. It answers what led Hubbell to work on his 2001 book. According to this narrative, the neutral theory comes out of the statistical project of describing the patterns and processes of relative species abundance distributions in communities using mathematical models.

Chapter 2 then takes up two additional historical questions, which lead to two additional, though entwined, narratives. I first ask what Hubbell worked on *before* the book. From this perspective, the neutral theory emerges out of the debate over whether communities are in taxonomic equilibrium, or whether communities respond to disruption by returning to their previous state or by going to a new state. I then ask why Hubbell called his theory "neutral". On the received history, the neutral theory is called *neutral* because of the surprising nature and mathematical importance of the formal assumption of ecological equivalence of all individuals. But chapter 2 reveals and emphasizes the connection of "neutral" to "null", and the methodology of testing

biological hypotheses against appropriate *null* hypotheses/ models, which has swept across biology since the 1970s.

Different epistemological questions and different answers to these questions can be prompted by different narratives of the origin and development of the theory. A different historical picture of the origins of the theory emerges when these narratives are considered in addition to Hubbell's origin story. And a different epistemic picture of the usefulness of the neutral theory follows from these different origins. In addition to enriching the historical origins of the theory then, these narratives prompt the epistemic questions taken up in chapters 3, 4, and 5 of this dissertation. The image of the neutral theory and the controversy surrounding it that I arrive at in chapter 5 is very different to the currently received one which I presented above and now elaborate.

3. Hubbell's Origin Story

In the preface to his 2001 book *The Unified Neutral Theory of Biodiversity and Biogeography*, Stephen Hubbell tells a story of how the neutral theory presented in the book came to be:

“This book is the outgrowth of a graduate seminar and a series of undergraduate lectures on biodiversity and biogeography at Princeton University. However, the germ of the ideas presented here can be traced by my thirty-year-old fascination with the origin and maintenance of high tree species diversity in tropical forests, an interest that launched several large-scale studies of tropical forest diversity and dynamics in the late 1970s. The earliest version of the present theory was published over twenty years ago in a paper in *Science* on tree diversity in a tropical dry forest in Costa Rica (Hubbell 1979).

At some point now lost to memory, while teaching the theory of island biogeography, I wondered

what would happen if a process of speciation were incorporated into the theory. I did not actually attempt this in the formal context of a mathematical theory until 1995.” (Hubbell 2001, xi)

The heart of this story is the question, “What would happen if MacArthur and Wilson’s Theory of Island Biogeography, a very prominent theory of in ecology, was augmented with a speciation process?” In another written version of the story he is more specific,

“I revisited neutral theory only in the mid-1990s, when a student in my Princeton biogeography class asked, why doesn’t the theory of island biogeography include a process of speciation, and what would happen if it did? I did not know, and I set about finding out.” (Hubbell 2008, 144c)

This story is interesting for a number of reasons. First, it perfectly fits a particular type of origin story about a scientific theory, wherein a stroke of insight or eureka moment leads a great theorist to unfold a great innovation. In the story of Newton’s apple, a version of which Newton himself is said to have told much later in his life, Newton is visiting home at Woolsthorpe Manor from Cambridge and hiding from the plague. He is sitting under an apple tree when he sees an apple fall, which may or may not hit him on the head, and in that instant comes to the law of universal gravitation.³ The story of how Darwin allegedly came to the theory of evolution by natural selection by looking at the differences between finch species on different Galapagos islands is another example, and they are common.

Second, it is refreshingly told by the chief actor themselves, as was Newton’s apple story. Third, the middle of Hubbell’s origin story suggests a deeper history (those lines in the quotation at the beginning of section 3, from “However” to “Hubbell 1979”), and read like they are being said so that they can be ignored. They seem to say, ‘while there a lot

³ Historians are divided about both the historical accuracy of these stories, and whether they are harmful myths. Newton’s apple story is generally accepted to have some historical accuracy, Darwin’s finches does not. On Newton, see (Christie 2016).

of earlier work, things really took off after this question was so clearly posed'. Newton's apple and Darwin's finches erase the (other, in the case of Newton) events that actually led to the theories of universal gravitation and evolution by natural selection from common descent. Fourth, the story reflects much more than how Hubbell himself sees the history he was shaping and being shaped by. It would not be very interesting to explain that Hubbell's story greatly oversimplifies and misleads the reader about the origin of the neutral theory of ecology. People are notoriously bad at remembering important events in a faithful way. The story is very interesting though if you imagine it resulting from Hubbell telling and retelling, framing and re-framing where the neutral theory came from.

I propose this story, Hubbell's 'Origin Story', as a jumping-off point for delving in to the origins of the neutral theory of ecology. Hubbell's origin story has been shaped not only by pressures external to Hubbell, but also by Hubbell's own ideas about scientific methodology and scientific innovation. This chapter and the next are concerned with digging into the origin story as *history*.⁴

3.1 A Shortcut from The Theory of Island Biogeography to The Unified Neutral Theory of Biodiversity and Biogeography

MacArthur and Wilson published the Theory of Island Biogeography (*TIB*) in a paper in 1963 (MacArthur and Wilson 1963) and book in 1967 (MacArthur and Wilson 1967). The theory is about biogeography in that it is concerned with *species richness*, the number of species on an island, and the *species-area relationship (SAR)*, the relationship between species richness and the size of that area. The TIB is used to predict and causally

⁴ The story could also fruitfully be analyzed as an origin story as *myth*.

explain SARs on islands. The SAR is one of the most straightforward patterns to measure and a number of interesting relationships have been observed. For example, an oceanic island will tend to have fewer species on it than a region of equal area on the mainland, and islands which are further away tend to have few species per area than islands closer to the mainland. Here are two graphical representations of actual SARs used by MacArthur and Wilson. Figure 1.2 comes from the 1963 paper.

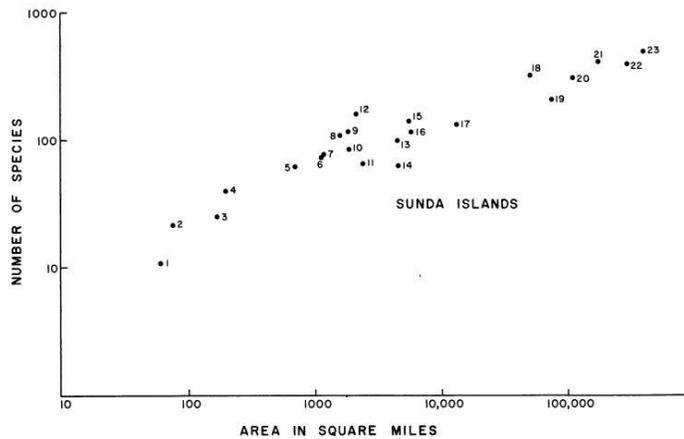


FIG. 1. The numbers of land and freshwater bird species on various islands of the Sunda group, together with the Philippines and New Guinea. The islands are grouped close to one another and to the Asian continent and Greater Sunda group, where most of the species live; and the distance effect is not apparent. (1) Christmas, (2) Bawean, (3) Engano, (4) Savu, (5) Simalur, (6) Alors, (7) Wetar, (8) Nias, (9) Lombok, (10) Billiton, (11) Mentawai, (12) Bali, (13) Sumba, (14) Bangka, (15) Flores, (16) Sumbawa, (17) Timor, (18) Java, (19) Celebes, (20) Philippines, (21) Sumatra, (22) Borneo, (23) New Guinea. Based on data from Delacour and Mayr (1946), Mayr (1940, 1944), Rensch (1936), and Stresemann (1934, 1939).

Figure 1.2 - Species-Area Plot of Sunda Islands

From (MacArthur and Wilson 1963, 374). Used with permission from John Wiley and Sons Inc.

Figure 1.3 comes from the 1967 book.

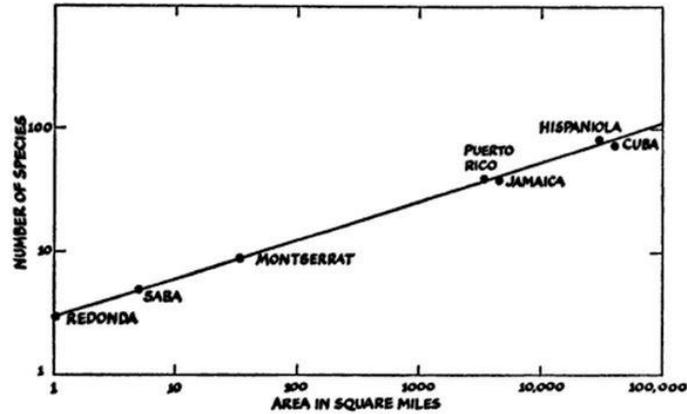


FIGURE 2. The area-species curve of the West Indian herpetofauna (amphibians plus reptiles).

Figure 1.3 – Species-Area curve of West Indian Herpetofauna

From (MacArthur and Wilson 1967, 8). Used with permission from Princeton University Press.

MacArthur and Wilson’s theory is that “the number of new species entering an island may be balanced by the number of species becoming extinct on the island.” (MacArthur and Wilson 1963, 374) This explains the species-area relationship analogously to how the water level in your sink is explained with the inflow rate from the faucet and the outflow rate of the drain. The theory is an (time-independent) *equilibrium* theory because it assumes that the number on species stays a constant.⁵ This is analogous to solving the sink problem when the water level is a constant.

⁵ It does not produce a *taxonomic* equilibrium, where the particular species in the community remains the same. I discuss the issue of taxonomic equilibrium in chapter 2, section 2.1.

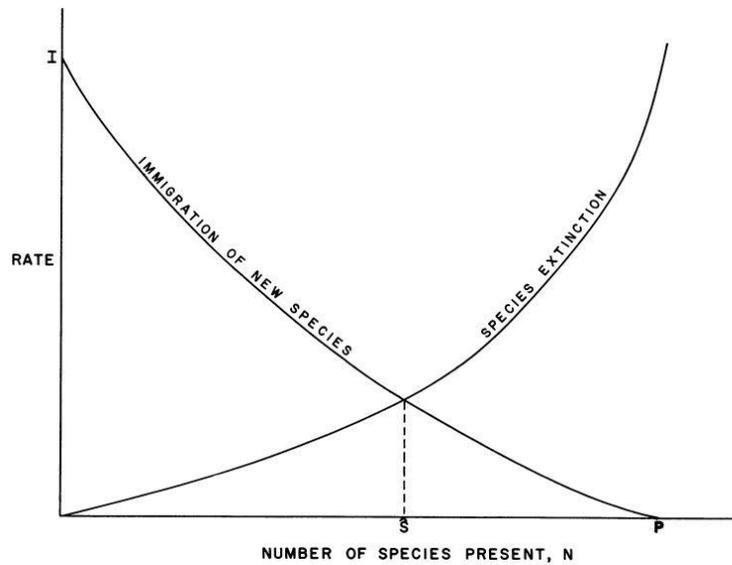


Figure 1.4 – Island Biogeography Model

From (MacArthur and Wilson 1963, 376). Used with permission from John Wiley and Sons.

They presented their theory through a graphical representation, as shown in figure 1.4 above. The downward-sloping curve represents the rate at which new species (not already on the island) immigrate to the island. The immigration rate curve I has negative slope because the chance of a new species immigrating is greatest when there are the fewest species present on the island and this chance decreases as more new species arrive. The immigration rate drops to zero once all of the species present in the source pool, P , are present on the island. The upward-sloping curve represents the rate at which species go extinct on the island. The extinction rate curve E has positive slope because the probability of more species going extinct scales with the total number of species, and because competition for space will reduce population sizes as species richness increases. While MacArthur and Wilson gave reasons for why actual curves will probably be concave, the theory is compatible with any monotonic curve.

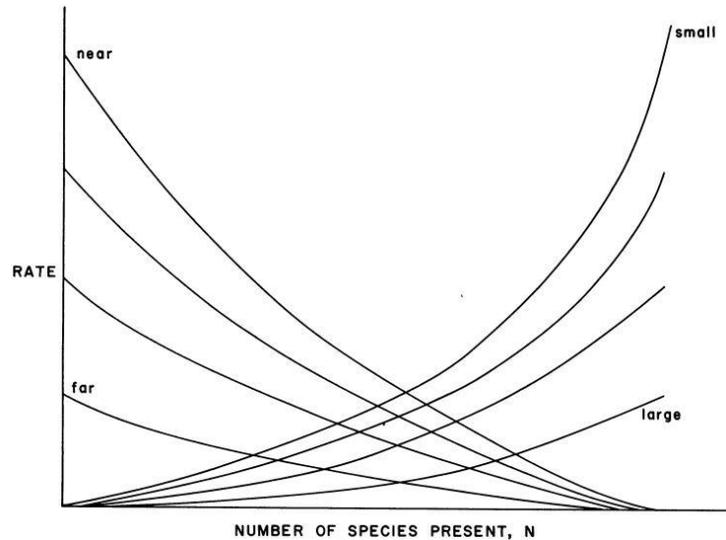


Figure 1.5 – Variation in Island Biogeography Models
 From (MacArthur and Wilson 1963, 377). Used with permission from John Wiley and Sons.

This simple theory explains differences in species richnesses N across islands in terms of differences in immigration and extinction rates, as the figure 1.5 above shows. It also explains differences in immigration extinction rates in terms of differences in distance to the island from the mainland source pool and differences in the size of the island. Extinction rates scale inversely with island area and immigration rates scale inversely with island distance. If two islands have the same size and are indiscernible except that one is further away, then their extinction rates will be the same but the island further away will have a lower immigration rate and so lower species abundance. Similarly, if two islands are the same distance away from a mainland and are indiscernible except that one is larger, then the immigration rates will be the same but the extinction rate will be higher on the smaller island and so it will have a lower abundance. These differences in abundance are neatly represented in figure 1.5. The distance effect is explained by through imagining two islands with the same size and extinction rate but with different distances and immigration rates. The theory of island biogeography has

proved to be empirically successful, fecund in terms of giving rise to analogous models of new domains, and well-respected within ecology and biology.

Hubbell's origin story explains the neutral theory of ecology as the outgrowth of the TIB. The origin story was meant to be both a conceptual presentation of the development of the neutral theory and a historical claim. Hubbell claimed that the neutral theory of ecology rectified two shortcomings of the theory of island biogeography (Hubbell 1997, 2001). First, island biogeography only explained species richness and not relative species abundance. Second, island biogeography did not include speciation (or a metacommunity). He has a conceptual presentation of the following form in mind.

To begin, assume that you want the TIB to represent SADs in addition to species richness and species-area relations. SADs contain strictly more information than species richnesses and so can be used to distinguish communities with the same species richness.

The TIB models species richness as the balance the number of new species immigrating to a community and the number of species going extinct in the community. In this way the TIB is based on *dispersal* and *demographic* processes. The TIB additionally assumes that each species is equally likely to immigrate or go extinct; immigration and extinction rates are neither dependent upon species nor on current abundances. In this way the TIB is *neutral* at the species level. The main difference required to model SADs is moving from the species level to the individual level. Dispersal and demographic processes acting on individuals can either leave their current relative abundance the same, increase it, or decrease it. Neutrality can be extended to the individual level by assuming that the rates at which a species increase, decrease, or

maintain their current abundances is equal for all species and independent of current abundances.

By moving to the individual level, drift must be included as an additional process to immigration. This is because births and deaths of individuals must be tracked. As I described above in section 1, extinction is then no longer an independent process, but a consequence of drift. Now Hubbell's second rectification comes in – adding speciation. The reason that speciation must be added is that drift and immigration (with saturation and neutrality) alone do not yield any time-independent equilibrium SADs that have more than one species. But, Hubbell's second insight was to see that speciation can be added to balance out the extinctions that drain the metacommunity and local community of diversity via drift and emigration. The SAD for a community is then solved for by defining functions for these rates and assuming that the SAD does not depend upon time and so arrives at an equilibrium.

In this way, by rectifying two shortcomings of the theory of island biogeography - its failure to describe relative species abundance distributions and its failure to include a speciation mechanism - and doing a lot of mathematics, one can arrive at the neutral theory of ecology. Now I hope you will agree, this is an effective way of presenting the neutral theory to an audience who knows the TIB. It really seems as if someone could have been struck by an insightful question and proceed to unfold out this theory. Here, the innovation of the theory stems not in any failed prediction or new information. It stemmed rather from specific aims for doing more with a useful but limited theoretical framework.

4. The origin story as history

In this section I consider Hubbell's origin story as a historical claim about the origin of the neutral theory of ecology. I believe that the story also serves as a myth about scientific innovation and methodology. But it is primarily intended as a historical explanation of where the neutral theory presented in the book came from.

I first show why the story is not historically accurate. Then in section 5, I expand upon the narrative of the origin story by investigating where the neutral theory presented in the book came from. Chapter 2 picks up here and investigates the origins of the neutral theory by asking two additional questions. My point is not to prove Hubbell wrong about the history of his own theory. My point is twofold: first, to show that the neutral theory of ecology is interesting as more than a formal model of patterns. Second, to show that regarding any one narrative of the neutral theory as the only narrative leads to an impoverished and misleading epistemic view of the usefulness and success of the theory. This second point will be taken up in later chapters of my dissertation.

But it is actually much more complicated than that

Hubbell's origin story misrepresents the actual historical development of the neutral theory in two significant ways. First, Hubbell was not adding speciation to the theory of island biogeography, he was adding it to his own *community drift model* of biodiversity. In Hubbell (Hubbell 1979), the community drift hypothesis has already produced a theory of biodiversity and SADs in the style of the theory of island biogeography. The TIB was only a model of species richness and richness per area, but

new patterns in the data such as the observation that most species in a community are rare made a theory of SADs attractive. Hubbell alludes to the community drift model in what I am calling Hubbell's origin story, as quoted in section 3, from the book. But within Hubbell's origin story, the student's question washed out the influence of the community drift model on the neutral theory, instead connecting the neutral theory to the theory of island biogeography. But, I show below, if Hubbell had not worked on the community drift model, he probably would not have created the neutral theory. And if the student had not asked Hubbell about adding speciation to the TIB, the neutral theory could still have come about. This is because of the second misrepresentation.

Second, many ecologists had observed the possibility and value of adding speciation to the mechanisms increasing diversity in a biogeography model. In his 1974 review of the TIB, Daniel Simberloff, a student of Wilson's (who got his PhD the same year (1969) as Hubbell), explained how the scale on which one is focusing determines which forces are balancing each other. On larger spatial and temporal scales, speciation balances with extinction, while on shorter spatial and temporal scales, immigration balances with extinction (Simberloff 1974, 177). Simberloff was part of the Woods Hole Group who developed the MBL model of clade diversity in paleobiology. The MBL model was explicitly constructed to work like the theory of island biogeography, except to balance speciation and extinction events.

Hubbell was aware of all of this.⁶ In 1984, Hubbell claimed that, "... the ultimate limit to species richness in local tropical rainforest tree communities is set by large regional scale speciation and extinction rates in conjunction with biogeographic patterns of, and limitations to, species dispersal." (Hubbell 1984, 8) He goes on later to make the 'conceptual incompleteness' claim about the theory of island biogeography theory,

"...addressing the question of origin, what are the spatial and temporal scales of speciation and extinction? By itself, community ecology is incapable of answering both the origin and maintenance questions about tropical forest diversity. The ecological theory of island biogeography is conceptually incomplete in this regard. No account is given of the origin of the species diversity in the "source area", which in the theoretical construct, ultimately sets the diversity in habitat isolates? Nor is any provision made for the origin of endemics in these habitat isolates. Thus, we clearly need a macroevolutionary theory of community ecology for tropical forest diversity - indeed for communities at all latitudes. After all, it is still unclear the extent to which there are any principles of speciation and forest ecology which are unique to the tropics, in spite of many speculations to this effect." (Hubbell 1984, 10)

By 1984, Hubbell had already constructed and began using the community drift model, based on drift and immigration. This paper further shows that Hubbell already believed that a full account of diversity must include drift, immigration, and speciation. It also shows that the idea that the theory of island biogeography needed to be augmented with speciation, to remedy its inability to account for where diversity comes to a community

⁶ Digging into the connection between the MBL model, clearly a neutral model, only sheds more darkness on the role the student's question played for Hubbell. In Ch. 8 of the book, Hubbell cites (Levinton 1979) as a presager of the neutral theory of ecology because of his balancing speciation and extinction with a verbal model of phylogenies. This paper is cited by Hubbell and Foster 1986 briefly and so it is not only in hindsight that Hubbell sees this connection. Yet why was this not impetus enough, Levinton drawing on equilibrium models like island biogeography, to think to add speciation into the community drift model, already based on the theory of island biogeography? I don't

from, was in the air. Additionally, as I show in chapter 2, Hubbell did not base his community drift model on the theory of island biogeography. Because all of the important pieces of Hubbell's origin story - drift, immigration, and speciation - are found together by 1984, the role of the student's question in Hubbell's origin story should lose much of its novelty and insight.

Hubbell's origin story considered as history could still explain why Hubbell returned to his earlier community drift hypothesis in 1995. It could explain why it was *then* that Hubbell put everything together, even if it downplays the role of the community drift model. Surely the presence of all the pieces alone does not necessitate putting them together. But Hubbell did not add speciation to the TIB to get neutral theory, rather he added speciation to the community drift model. Therefore, assuming that the student in his seminar did ask the question and did spur him to work again, Hubbell's origin story remains a biased picture of the historical origins of theory. In section 5 of this chapter and in chapter 2, I present three different narratives which answer three questions regarding the origins of the neutral theory of ecology. These fill in and contextualize Hubbell's development of the neutral theory of the book.

5. Where did the Unified Neutral Theory come from?

Formal models of relative species abundance

distributions

know what differences happened from 1986 to 1995 or if it was just a case of needing to hear something so many times before the light goes off.

Three longstanding projects in community ecology (often now called macroecology⁷ in this context) are to (1) mathematically describe macroecological patterns of abundance and diversity across communities, (2) predict how macroecological patterns vary in time and space, and (3) give causal explanations of these patterns. Two patterns that have received sustained attention are the species-area relationship (SAR) and the relative species abundance distribution (SAD).⁸ These projects are connected in complicated ways because description, prediction, and causal explanation are connected in complicated ways.⁹ Historically, description has been worked on for the purposes of measurement and comparison of diversity of communities in space and time, as well as for prediction and explanation. Explanation has been dependent upon description and prediction, but not exhausted by them.

The strategy taken by the proponents of the *neutral theory of ecology* has been to construct a mathematical model on the basis of ecological principles which predicts a statistical distribution, use the *neutral model* to quantitatively describe and predict SADs very accurately, and regard accurate descriptions and predictions (here after “model-fits” unless otherwise specified) as reason to give causal explanations of SADs and other patterns using the neutral theory.

⁷ This term was coined in (Brown and Maurer 1989).

⁸ Paleobiological patterns such as how species richness and species abundances vary in the fossil record is another pattern of interest.

⁹ I have Levins (Levins 1966) in mind here. Generality, precision, and realism are three different goals or virtues of scientific models. The generality of a model is the number of situations which a model can apply to. The precision of a model is the mathematical precision of a model. Qualitative models are less precise than quantitative models, for example. The realism of a model is how well the model takes into account the complexity of the target system. A given modeling strategy usually picks some goals and ignores others. Of the modeling strategies Levin’s describes, Hubbell’s is closest to *Sacrifice realism to generality and precision*.

This section presents the neutral theory of ecology in the context of the projects of using formal models to describe, predict, and explain patterns. I proceed in four steps which follow the four layers involved in proceeding from data collection to explanation. Section 5.1 explains the abundance data collected, focusing on relative abundance data. Section 5.2 explains how representations of relative abundance data as SADs are used to extract patterns from abundance data. Section 5.3 explains the historical development of the statistical distributions as representations of relative abundance data used to describe and predict the patterns in abundance data. And section 5.4 explains the historical development of causal and other interpretations of statistical distributions of representations of relative abundance data used to give (and undermine) causal explanations. This is the expanded story that Hubbell has in mind when telling his origin story and this narrative dominates discussion of the neutral theory. And this way of telling the origin of the neutral theory sets up a particular way of viewing the development of and controversy surrounding the neutral theory. The next chapter in turn complicates the origin story with other narratives, which I then use to provide another way of viewing the development of and controversy surrounding the neutral theory.

5.1 Abundance Data

Data about the abundance and diversity of ecological communities is collected by sampling. In this narrative I only focus on the patterns on SADs. The situation for SARs has the same epistemic structure, only the data is different. Ecologists studying communities often limit their study to *single-trophic level communities*. A trophic level is

roughly all the organisms which eat the same kinds of thing and are eaten by the same kinds of thing. Plants/ primary producers are a trophic level. Herbivores/ primary consumers are a trophic level. And so on. Trophic levels exclude interactions such as predation. In a sample of a bird community, raptors will not be included with songbirds. From here on, I refer to single-trophic level communities as “communities” unless otherwise specified.

A complete census of a (single trophic-level) community will minimally include the species of every individual in the community. This can be organized in a table showing the number of individuals in each species counted. Most communities cannot be censused completely. It takes a lot of energy, but a tract of forest can be moved through systematically quickly enough that the forest is mostly the same as when the sampling began. But insect communities reproduce much too quickly to count them all, and they move around.¹⁰ Bird communities fall in between. A complete census is ideal, but typically ecologists settle for a limited sample of the individuals in a community. The important issue here is that the sample be random with respect to commonness and rarity (Preston 1948, 254). For example, moths can be sampled with light traps and though this is not random with respect to how species respond to light, it does not care if the individual caught is of a rare or common species. The collections of naturalists in museums are not random samples, because naturalists are drawn to rare and unsampled species.

¹⁰ Microbial communities are obviously even more extreme. Microbial ecologists are very distinct from *macrobiological* ecologists.

5.2 Representations of Relative Species Abundance

A first step towards understanding the patterns of commonness and rarity in a way that permits cross-community comparison is representing the data across communities in the same way. Because different communities contain different species, in practice ecologists interested in the commonness and rarity of species across communities leave out the information on the identity of the species. They also often leave out any species which have zero members, as might exist at one time when surveying the same community over time. The information of a sample will include only the number of species and the number of individuals in each species then present. This information is represented graphically several ways. Two of the most common ways to present this information are Preston plots and Whittaker plots.¹¹

In order to represent the idea that species are not common or rare absolutely, but common or rare relative to other species in the same community, Preston (Preston 1948) represented species abundances data in the following way. Create a series of bins A, B, C, ... N, where bin B includes the species with twice as many members as bin A, C twice that as bin B, and so on. Each bin holds a *doubling class* or *octave*. In practice, bin A held the species with ≤ 2 members, bin B held species with 3-4 members, C held 5-8, and so on. Preston chose doubling classes, but this technique works the same for base 3 or base 10. Figure 1.6 shows data plotted in for three different choices of bases.

¹¹ These can be regarded as *models of data* following (Suppes 1962).

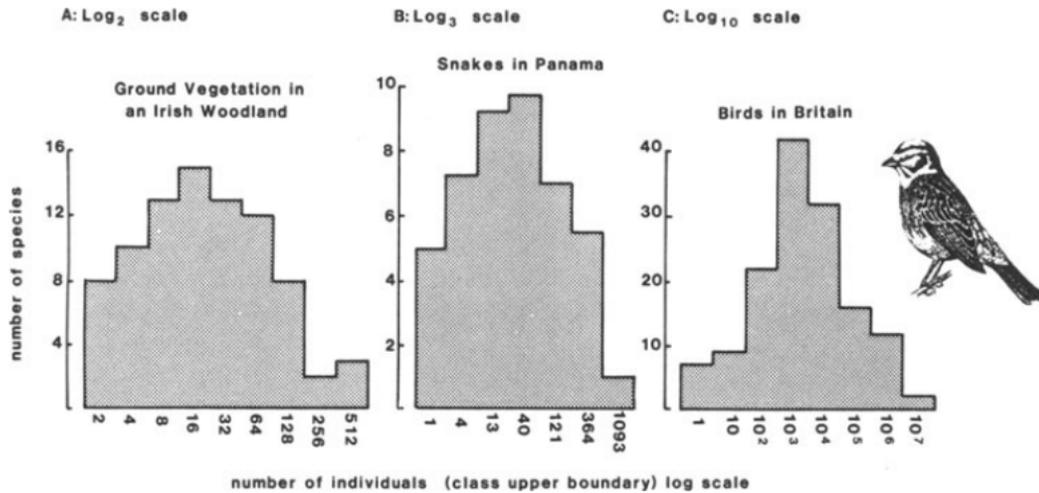


Figure 2.7 The log normal distribution I. The ‘normal’, symmetrical bell-shaped curve is achieved by logging the species abundances on the x-axis. A variety of log bases can be used. (A) \log_2 . This usage follows Preston (1948). Species abundances are expressed in terms of doublings of numbers of individuals. For example successive classes would be 2 or fewer individuals, 3–4 individuals, 5–8 individuals, 9–16 individuals, 17–32 individuals and so on. It is conventional to call the classes, octaves. The graph shows the diversity of ground vegetation in a natural deciduous woodland at Banagher in N. Ireland (see Figure 4.2 and Chapter 4). (B) \log_3 . Instead of doublings the successive classes refer to treblings of numbers of individuals. Thus in this example showing the diversity of snakes in Panama (data from Williams, 1964) the upper bounds of the classes are 1, 4, 13, 40, 121, 364 and 1093 individuals. Although used widely by Williams (1964) \log_3 is rarely employed today. (C) \log_{10} . Classes in \log_{10} represent increases in order of magnitude 1, 10, 100, 1000, 10000, 100000. This choice of log base is most appropriate for very large data sets, as for example in this case the diversity of birds in Britain (data from Williams, 1964). In all cases the y-axis shows the number of species per class.

Figure 1.6 – Lognormal Distribution in Different Bases

From (Magurran 1988, 20) Used with permission from Springer Science and Bus Media B V.

The x-axis records the abundance classes and the y-axis records the number of species in the abundance class. A part of the line is read as “there are y many species with abundance-class-x-many members”.

Preston plots show the *relative* abundances of species well, but do so at the cost of reducing the amount of information in the sample. Knowing the abundance class of a species does not tell you the exact abundance of that species. Other ways of representing the data include all the information except species identity.

A natural way to represent commonness and rarity is in a pie chart. Each slice of the pie represents a different species and the size of the slice represents the relative abundance of that species. A Whittaker plot includes the same complete set of information as the pie chart, but does so as a line on a graph. Figure 1.7 (from (Hubbell 2001, 4)) shows Whittaker plots of five different communities.

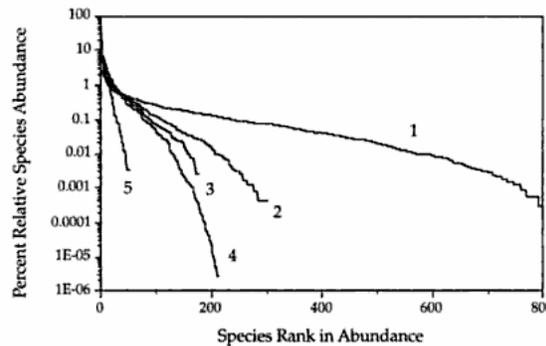


FIG. 1.1. Patterns of relative species abundance in a diverse array of ecological communities. Species in each community are ranked in percentage relative abundance from commonest (*left*) to rarest (*right*). The percentage relative abundance is log transformed on the *y*-axis. 1: Tropical wet forest in Amazonia. 2: Tropical dry deciduous forest in Costa Rica. 3: Marine planktonic copepod community from the North Pacific gyre. 4: Terrestrial breeding birds of Britian. 5: Tropical bat community from Panama.

Figure 1.7 – Comparison of Species Abundance Distributions
From (Hubbell 2001, 4). Used with permission from Princeton University Press.

The x-axis represents the rank in abundance of a species and the y-axis represents the percent relative abundance, often plotted on a log₁₀ scale. A point on a line is read as “the xth ranked species makes up y% of the community”.

Whittaker plots permit easy visual comparison of the *evenness* of a community. A community is even to the degree that species have similar relative abundances to each other, and this corresponds to a line being more horizontal. For example, these curves show that community 1 is more even than any other.

SADs show one of the few empirical generalizations in ecology which seem to be universal - communities have few very abundant species and many rare species.¹² To standardize terminology, I refer to a representation of relative species abundance such as a Preston plot or Whittaker plot as a SAD.¹³ In order to permit more detailed analysis of and comparisons between SADs, more than visual inspection is required. This leads to the use of statistical distributions of SADs.

5.3 Statistical distributions of representations of relative species abundance

The general goals being addressed with formal distributions¹⁴ of SADs are describing, predicting, and explaining SADs and other patterns. In this section I show some of the major developments in statistical distributions of SADs used for description and prediction. I then draw some conclusions for the epistemic problem of inferring process from pattern, which is the subject of the next subsection. I present the customary, important developments - Fisher's logseries distribution, Preston's lognormal distribution, and Motomura's geometric distribution.

Fitting statistical distributions to SADs took off with the 1943 paper by R. A.

¹² (McGill et al. 2007) calls this one of ecologies few true universal laws. This is a *law of phenomena* (contrasted with laws of causes) in Whewell's terminology, and *distribution* (contrasted with causal regularity) in Waters' (Waters 1998) terminology.

¹³ There are other ways to represent SADs. For a survey see (McGill et al. 2007) and (Magurran 2004, Ch 2)

Fisher, A. Steven Corbet, and C. B. Williams (Fisher, Corbet, and Williams 1943). In this paper, Fisher proposed that SADs based on censuses would fit the *logseries distribution* well. This meant that the logseries distribution could be used to estimate the true, complete SAD of a community from a random sample because the SAD of a random sample has the same statistical properties as the true SAD, just with more data points. For all statistical distributions of SADs, the goal is to describe a SAD using a statistical distribution that fits the SAD very well. When ecologists claim that “this sample displays a logseries (or other) distribution” what they mean is “this sample displays a statistical distribution that cannot be distinguished from a logseries in practice” (Pueyo 2006, 399).

Fisher developed the logseries distribution for this paper. For our purposes, the important characteristics of the logseries distribution are the following (Magurran 2004, Ch 2). The logseries states that the number of species with 1, 2, 3, ... n members is:

$$\alpha x, \alpha x^2 / 2, \alpha x^3 / 3, \dots, \alpha x^n / n$$

where $0 < x < 1$.

Because each term in the series is greater than the next, applying the logseries to SADs predicts that there will be more species with 1 member than 2, more with 2 than 3, and so on. The logseries further relates N , the total number of individuals, to S , the total number of species. Values of α and x are estimated from the data set itself by fitting the series to data based on the following relationships:

¹⁴ I am choosing my terminology to standardize the reasoning patterns being followed. “Statistical distributions” here are also called “statistical models”. But I will use the term “model” to refer to the combination of distribution produced *and* the interpretation of the distribution. Statistical models produce statistical distributions which have statistical interpretations, while biological models produce statistical distributions which have biological interpretations. But statistical distributions and models are compared to representations of data (data models) with “model-fits”.

$$S/N = (1 - x/x)(-\ln(1 - x))$$

$$S = \alpha \ln(1 + N/\alpha)$$

Empirically, x is almost always >0.9 and never >1.0 . If $x \approx 1$, then α measures the number of species expected to have 1 member. For this reason, α is still used today as a way of comparing the diversity of different communities.

This paper showed that the logseries distribution developed by Fisher produced good model-fits to the Lepidoptera abundance data gathered by Corbet, and suggested that it would do so for other data. But the authors did not propose any biological explanation for why this was, such as giving a biological interpretation of the logseries distribution. Their project was descriptive and predictive.

In 1948, Preston argued that the good fits obtained between the logseries and SADs was an artifact of non-random and too limited sampling (Preston 1948). Preston proposed the *canonical lognormal distribution* as a better statistical distribution for fitting SADs. The lognormal distribution describes the result of the log-binned abundance classes following the normal distribution.

Thinking in terms of a Preston plot, the lognormal distribution states the following:

$$S(R) = S_0 e^{-a^2 R^2}$$

Where $S(R)$ = the number of species in the R th octave from the mode ($+R$ to the right, $-R$ to the left); S_0 = the number of species in the modal octave (the octave which is the mode of the curve); and $a = (2\sigma^2)^{-1/2}$ (where σ is variance). a is a constant fitted to the data and controls the inverse width of the curve (Magurran 2004, 32, Preston 1948, 257). The

parameters a and S_0 can be determined from knowing the total number of species and the total number of individuals in the sample. Empirically, a is usually found to be around 0.2 (Preston 1948, 1962). The *canonical* lognormal distribution predicts that the largest class will be the mode, and that the rarest and most abundant classes will be equally (un)common. In contrast, the logseries distribution predicts that rare species will be much more common than the most abundant species.

Preston argued that this is what is found empirically in very thoroughly sampled communities. However, many samples of communities empirically fail to be lognormally distributed because of the difficulty of sampling all the rare species. Preston explained this as a sampling effect - if a perfectly random sample of a community would find a small fraction of an individual, then it is most likely that the individual will not be in the sample (Preston 1948, 262). Sampling will then underestimate rare species, and this will make the number of species with 1 member larger than the number of species with the most members. Preston saw that sampling truncates the lognormal distribution by cutting off some proportion of the left-hand side. The line truncating the lognormal is called the *veil line*, to the left of which is supposed to contain the missing octaves. Preston argued, supported by empirical sampling data that, as more random data is collected, the sample better approximates a lognormal distribution in the sense that the veil line is pushed further to the right. Over time, a distinct hump shape emerges showing that the number of species with 1 member is no longer more represented than the number of species with the most members.

Here are two figures Preston used to convey the idea. Figure 1.8 shows data compared with the best-fit lognormal distribution where the data appears to show a hump.

Here only the hypothetical left tail is behind the veil line. Remember that the far left of the curve on a Preston-plot represents the number of the rarest species.

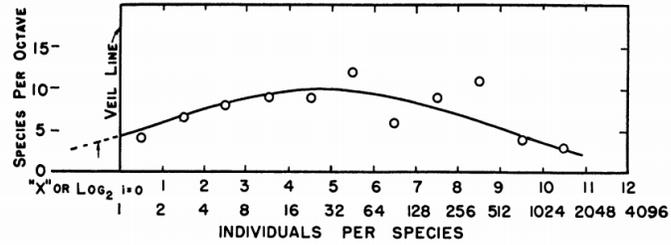


FIG. 2. Saunders' breeding birds. The octaves are definitely not equally filled, and the curve looks as if it is humped up in the middle.

Figure 1.8 – Lognormal Distribution fit to data
From (Preston 1948, 258). Used with permission from John Wiley and Sons Inc.

Figure 1.9 shows data that do not show the hump.

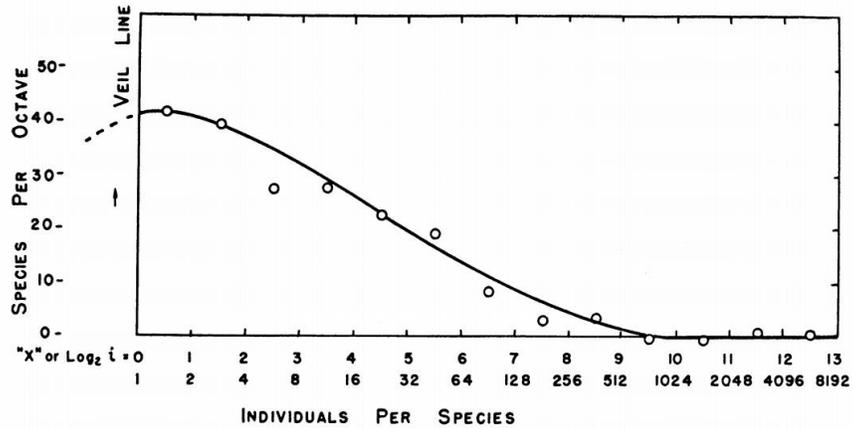


FIG. 4. Dirks' female moths. The hump seems to be very close to the veil. In Sweadner's moths of Wyoming (not illustrated), the veil has moved to the right of the hump, and all we see is the descending side.

Figure 1.9 – Lognormal Distribution fit to data with veil line
From (Preston 1948, 261). Used with permission from John Wiley and Sons Inc.

Here the hypothetical hump is centered on the veil line. By adjusting the veil line, truncated lognormal distributions can be made similar to logseries distributions (Magurran 2004, 40).

Logseries and lognormal distributions are less distinct visually in Whittaker plots

than in Preston plots. Figure 1.10 from compares Preston and Whittaker plots for the logseries (open triangles), lognormal (black dots), and power law¹⁵ (open squares) distributions.

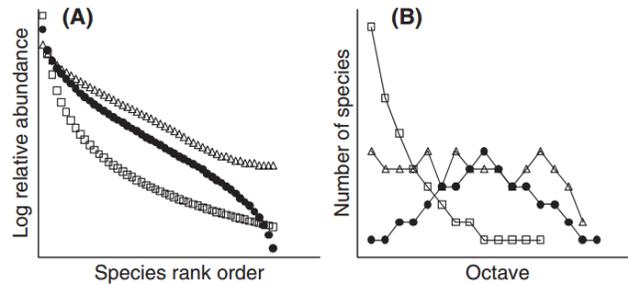


Figure 1. Basic shapes of species–abundance distributions in Whittaker (A) and Preston plots (B). Black dots: lognormal shape; open triangles: logseries; open squares: power law.

Figure 1.10 – Comparison of Whittaker and Preston plots
 From (Ulrich, Ollik, and Ugland 2010, 2). Used with permission from John Wiley and Sons.

Notice that the in Whittaker plot on the left, the right side of the curves differ significantly, where the lognormal is much less than the logseries. This shows that a lognormal-distributed community is less evenly distributed than a logseries-distributed community, because, Preston hypothesized, the data from the logseries-distributed community is missing the rarest species.

The logseries and lognormal remain the most widely used statistical distributions, but they are far from the only ones. Historically, the West learned in 1965 that Motomura had worked on the same issue and had proposed the *geometric distribution* in 1932. The geometric distribution is arrived at by sequentially dividing up a unit length in breaks of size k . The number of species with rank abundances 1, 2, 3, ... n have relative

¹⁵ The power law is another statistical model used to describe SADs. See (Ulrich, Ollik, and Ugland 2010) for a description.

abundances:

$$S(1, 2, 3, \dots, n) = k, k(1-k), k(1-k)^2, \dots, k(1-k)^{n-1}$$

where k is a constant that represents the relative proportion of the divisions (Magurran 2004, 48).

The geometric distribution produces the less evenness than the lognormal or logseries distributions. Motomura too proposed no biological interpretation of the geometric series.¹⁶

Figure 1.11 shows the characteristic shapes, especially evenness, of the geometric, logseries, and lognormal distributions on a Whittaker plot.

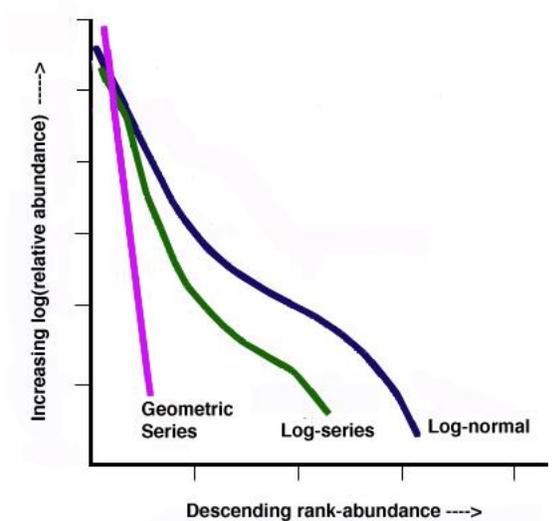


Figure 1.11 – Comparison of Geometric, Log-series, and Log-normal Distributions
Created by Aedrake09. Retrieved from
https://commons.wikimedia.org/wiki/File:Common_descriptiveWhittaker.jpg

Figure 1.12 shows another way these distributions are related to each other, with respect to the spatial scale of the community sampled (from)

¹⁶ An English translation of Motomura 1932 can be found as an appendix to (Mori 2013).

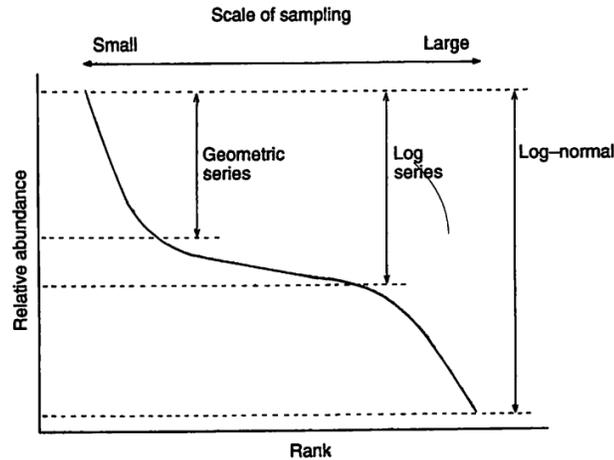


Fig. 21. Nested relationship between the geometric-series, the log-series and the log-normal patterns based on the scale of sampling.

Figure 1.12 – Relationship between Geometric, Log-series, and Log-normal Distributions From (Tokeshi 1993, 166). Used with permission from Elsevier.

Compare figures 1.11 and 1.12 with figure 1.6 above of the different SADs for communities and you can see that different SADs will be best fit by different distributions.

From this brief survey of the development of statistical distributions of SADs, several points about the epistemic issues raised by describing and predicting SADs. These will be relevant to the epistemic issues of explaining SADs discussed in the next subsection.

First, for the goal of description, it may be that the logseries and lognormal are able to give equally good model-fits to SADs. Additionally, one statistical distribution may fit the Preston plotting of the abundance data better than the Whittaker plotting of the abundance data.

Second, for the goal of prediction, if the logseries distribution gives the best model-fits to a SAD, then one can also test predictions about whether the logseries will continue to give the best model-fits to data from the same community which is more

extensive or more random. This may seem like an odd form of prediction, because no parameter values are being measured and then used to deduce distributions which can then be confronted with data. It might be called extrapolation, but it is a statistical sense of prediction that is used by ecologists. In particular, this kind of testing is useful for discriminating between two or more distributions which fit some limited data equally-well. More data can be used to discriminate between previously indiscriminable data.¹⁷ The distributions make different predictions about properties of communities such as (i) the number of species with the most members and (ii) the evenness of the SAD. Remember also that the statistical distributions discussed so far have been purely phenomenological and have no biological or other interpretation, and so are not meant to be, and in a sense cannot be, measured independently. What can be done is compare the parameter values which are arrived at by a process such as regression or maximum-likelihood techniques, which yield the values that give the best-fit of a distribution to data, across communities in time and space.

Third, for the goals of description and prediction, it can be difficult to find a crucial experiment to distinguish between two statistical distributions of SADs because of the flexibility allowed in their usage. A community may have better model-fits to the logseries than to the lognormal distribution. But this can be remedied by truncating the lognormal distribution and justified through the limitations of the sampling data. For limited samples, it is not always clear if a more extensive or random sampling would change which statistical distribution best fits the community.

¹⁷ This is the situation that the AIC model selection criteria is used for, and can be tested for. See especially (Forster and Sober 1994) for an introduction. I discuss the AIC in chapter 3.

5.4 Interpretations of statistical distributions of representations of relative species abundance

In 1957 (MacArthur 1957) and the 1960 (MacArthur 1960), Robert MacArthur expanded the ecological project from statistical description and prediction to explanation by proposing biological interpretations of statistical distributions of SADs:

“Earlier investigations, discussed elsewhere, fit known statistical curves of uncertain biological meaning to the data. A more fruitful approach seems to be to predict curves on the basis of simple biological hypotheses and to compare these with the data. Three such hypotheses will be discussed here.” (MacArthur 1957, 293)

MacArthur proposed three different ecological hypotheses for the abundance and diversity of communities.¹⁸ MacArthur assumed that each species occupies a niche in a community and conceived the three hypotheses as three ways that niches fit together. The minimal concept of a *niche* needed here is the set conditions of ecological resources used by a species.¹⁹ Resource use is then proportional to abundance. Following MacArthur, models based on the concept of a niche are called *niche apportionment models*.²⁰

MacArthur’s first hypothesis, “Nonoverlapping Niches” but now universally known as the “broken-stick model”, treats the environment as a finite stick and imagines

¹⁸ (Mori 2013) corrected two, incompatible, false beliefs about this project. First, that MacArthur was the first to propose a biological explanation of SADs. Second, that Motomura proposed a niche apportionment explanation (explained below) of his geometric distribution before MacArthur.

¹⁹ A niche also includes the way that an organism effects its environment. Some speak of a niche as the role played by a species or how it lives. Magurran (Magurran 2004) says that Tokeshi (Tokeshi 1993) argued that the literature here was confused between two concepts of a niche introduced by Hutchinson (Hutchinson 1957). The *fundamental niche* is the way an organism can live. The realized niche is the way an organism actually lives.

breaking the stick up into n segments by breaking the stick instantaneously at $n-1$ places. Intuitively, the length of each segment corresponds to the abundance of some species. The probability of a segment being broken is proportional to its length. The second hypothesis, “Overlapping Niches”, treats the environment as a stick again, but this time the size of one niche is determined by the distance between two points thrown at random on the stick and the overall distribution determined by n such throws. The third hypothesis, “Niches Particulate, Not Continuous”, the environment is made up of distinct urns into which individuals are places with random, independent turns.

Each of these hypotheses intuitively reflects a way that species might fit together to construct the community. From each of these hypotheses MacArthur derived a formula for the expected abundance of the r -th rarest species, and from this a statistical distribution. These three statistical distributions could then be compared on their model-fits to some sampled data of warblers in different communities. And from this descriptive and predictive problem, he could gain evidence for biological hypotheses to explain the distributions of abundance. MacArthur took the fact that the broken-stick distribution fit the sampled data consistently better than the other distributions as evidence for the non-overlapping niche ecological hypothesis. The idea that different species compete for space through differentially utilizing resources was tied to the view that no two species could utilize the same niche indefinitely and known as the *competitive exclusion principle*. MacArthur took his warbler abundance data as one form of evidence for competitive exclusion structuring warbler communities.

On the broken-stick model, the statistical distribution expected gives the number

²⁰ This term is due to (Sugihara 1980).

of species with rank abundance n , $S(n)$ has relative abundances:

$$S(n) = (S(S-1)/N)(1-n/N)^{S-2}$$

where S = total number of species and N = the total number of individuals

(Magurran 2004, 50).

The broken-stick model produces a distribution which is more even (horizontal) than the lognormal distribution. But for small sample sizes it is difficult to distinguish from the lognormal. Figure 1.13 compares the distributions on a Whittaker plot:

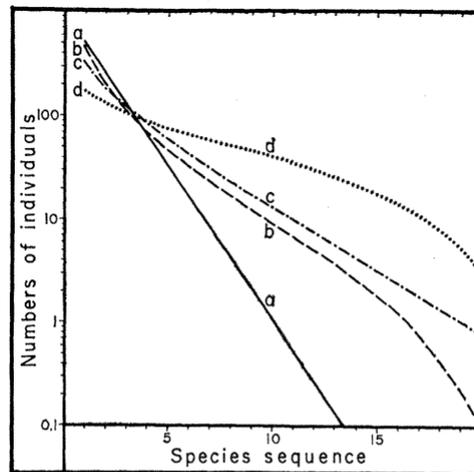


Fig. 1. Curves to fit dominance-diversity relations—four major hypotheses: (curve *a*) geometric series of Motomura (1), $c = 0.5$; (curve *b*) lognormal distribution of Preston (6), $a = 0.2$, $n_0 = 2.26$; (curve *c*) logarithmic series of Fisher, Corbet, and Williams (4), $\alpha = 3.542$; (curve *d*) random niche hypothesis of MacArthur (10). Numbers of individuals in the species, on the ordinate, are plotted against species number in the sequence of species from most to least abundant, on the abscissa. The curves are all computed for a hypothetical sample of 1000 individuals in 20 species.

Figure 1.13 – Broken-Stick Model on Whittaker Plot
From (Whittaker 1965, 251). Used with permission from The American Association for the Advancement of Science.

MacArthur's broken-stick model introduces two helpful distinctions going

forward. First, the broken-stick model is a *stochastic* model, while the logseries, lognormal, and geometric distributions are deterministic models. A deterministic model produces only one distribution given values of inputs. If there are S-many species and N-many individuals in a community, a deterministic model says that there will determinately be A many species in the smallest abundance class. A stochastic model specifies instead probabilities for the number of species in the smallest abundance class.

Second, the broken-stick model is a biological model and a niche apportionment model. A *biological model* produces a distribution on the basis of a biological hypothesis. Biological models are contrasted with *statistical models*. A niche apportionment model is a type of biological model based on apportioning relative abundance in a community on the basis on apportioning niche size. One alternative to niche apportionment models are *neutral models*.

One use of biological models is to give *biological interpretations* of SADs. The hope here is that fits between biological model and SAD can be used to give evidence to biological hypotheses, which can then be used to explain the causes of SADs. Some ecologists have also given *statistical interpretations* of SADs, which carries with it no biological interpretation.

In what follows, the difference between *distributions* and *models*, as in, but not limited to, the distinction between *statistical distribution* and *statistical model*, is the following: models have interpretations, while distributions do not. Because of the distinction between statistical distributions of SADs and interpretations of statistical distributions of SADs, most commonly used statistical distributions have been given both biological and statistical interpretations. This means that the inference from statistical

distribution to interpretation is one-many, and will in general be underdetermined. Below I give examples of this, before coming to neutral interpretations of statistical distributions.

5.4.1 Niche apportionment interpretations

MacArthur's work started a research program.²¹ The ecological project addressed is to construct statistical distributions of SADS, as discussed in section 5.3, which fit SADS well, on the basis of an interpretation in terms of niche apportionment. All are different scenarios for how a series of species' niches might interact and how species utilize resources, where resource use is proportional to abundance. While MacArthur hoped to test the competitive exclusion principle via the broken-stick model in order to learn about how niches pack together, subsequent work on niche apportionment models have only developed models where niches are nonoverlapping and contiguous.

Whittaker presented Motomura's geometric distribution and gave it a niche apportionment interpretation (Whittaker 1965). Assume a community had resources N and the most successful species utilizes a fraction k of the resources and so occupies kN of the community. If the second most successful species occupies fraction k of the remained $(1-k)$, they will occupy $k(1-k)N$. And the n th most successful will occupy $k(1-k)^{n-1}N$. This is the geometric series.

Sugihara modified the broken-stick model to be a sequential apportioning of niche

²¹ MacArthur himself disowned the methodology rather quickly (MacArthur 1966). After someone corrected the formula for the broken-stick model, he wrote,

“Let us hope these comments do not draw additional attention to what is now an obsolete approach to community ecology, which should be allowed to die a natural death.”

space and abundance and showed that a lognormal distribution would result (Sugihara 1980). This idea was generalized by Tokeshi who discussed eleven, including MacArthur's three, different ways of sequentially breaking the stick each with a niche apportionment interpretation (Tokeshi 1990, 1993, 1996). I will present several in order to show that the relationship between statistical distribution and model is one-many.

Whittaker's version of Motomura's geometric distribution is the most uneven and only deterministic niche apportionment model. A more general, stochastic form of the geometric model is the *dominance preemption model*. Here the stick is broken in two at random and the smaller section is chosen to be broken next. But instead of maintaining the same fraction, the smaller segment can be broken in any proportion at any step. If you imagine subsequent runs of this and each of the following stochastic models, subsequent runs of a finite number of breaks will probably not produce the same SAD. Each model rather specifies how probable each SAD and different models make different SADs more or less probable.

The *random fraction model* is a further generalization. Here all segments have equal chance of being broken, and the size of the break is also random. This model produces the lognormal distribution. The *MacArthur fraction model* has the stick broken at a random place. The probability that a segment will be broken next turn is proportional to its length. This would be satisfied if invading species are more likely to invade the niche of a species with a larger niche than a smaller one. This reverses the idea behind the dominance preemption model, where smaller niches are invaded. Finally, the *dominance decay model* is the opposite of the dominance preemption model. Here the stick is broken at random and the larger segment is chosen and broken at random. This produces the

most even distribution. Figure 1.14 shows their relationship graphically on a Whittaker plot:

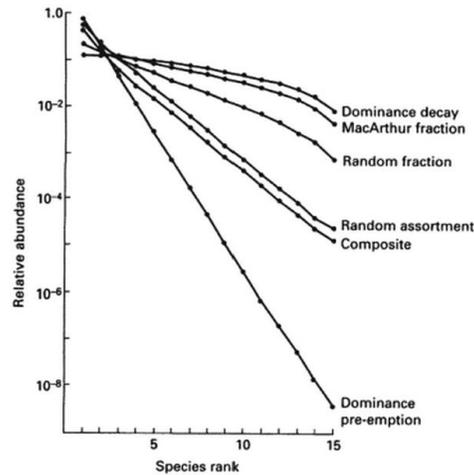


Fig. 7.3 Rank-abundance patterns of some niche apportionment models.

Figure 1.14 – Comparison of Niche Apportionment Model Distributions From (Tokeshi 1999, 190). Used with permission from John Wiley and Sons.

For Tokeshi, a given niche apportionment model explained a SAD when the model's assumptions about probabilities are satisfied by the community from which the SAD was drawn. But because determining which assumptions of a model are satisfied is difficult to do directly, one source of indirect support for the way that niches fit together in space and time in a community was through model-fit. The hope was to be able to gain evidence for the biological interpretation from the model-fit of the statistical distribution and the SAD.

5.4.2 Statistical interpretations

Cohen showed that the same statistical distribution was expected from both MacArthur's broken stick model and the *exponential model* (Cohen 1968). The exponential model assumes that the abundances of species are drawn randomly from an

exponential distribution. He argued that, if species abundances were exponentially distributed, which could be for either statistical or biological reasons, then the SAD should fit the exponential model. Cohen took his demonstration of an underdetermination problem to show that you cannot infer from good model-fits between the broken stick model and a SAD that the communities' niches are nonoverlapping and contiguous (Cohen 1968, 167). The same argument can be run for any of the statistical interpretations of any of the niche apportionment models discussed so far.²²

The most famous example of a statistical explanation is by May of the lognormal distribution (May 1975). May argued that many large data sets show a lognormal distribution because of the central limit theorem. The central limit theorem of probability theory states that the distribution of a sum of a large number of independent and identically distributed variables will probably be approximately normally distributed, no matter the underlying distribution of the variables.²³ And when the logarithm of a product of independent and identically distributed variables becomes approximately normally distributed, the product of these variables becomes log-normally distributed. For SADs, the variable is the number of individuals in a species, which is log transformed in both Preston and Whittaker plots (Magurran 2004, 35). If the ecological world is a mess and there are umpteen different processes effecting species abundances, then you should expect SADs be approximately lognormal. This is taken, again, to undermine any

²² (Gotelli and Graves 1996, 55) have a nice chart that tracks biological and statistical interpretations of all the standard statistical models of SADs.

²³ <http://www.math.uah.edu/stat/sample/CLT.html>

inference to biological processes from SADs.²⁴

Nekola and Brown generalized this type argument in an interesting way (Nekola and Brown 2007). They note the ecological patterns most often used to infer process from pattern - including SADs and SARs - have analogs across “physical and human economic, social, and artistic systems” (Nekola and Brown 2007, 189). Their examples for some patterns that are analogs of SADs are: “North American precipitation classes, stock volumes for all publicly traded US corporations, song performances for Cowboy Junkies [a band] setlists, and citation frequencies for scientific papers” (Nekola and Brown 2007, 190) They then argue that, because the analogs of SADs and other ecological patterns actually have the same range of shapes as SADs, any uniquely ecological explanation of SADs is unsatisfactory. They call for a general, universal explanation based on the statistical behavior of complex systems, not an explanation from ecological processes of the kind sought who try to get ecological processes from patterns.

5.4.3 *Neutral interpretations*

Another set of interpretations of SADs are the neutral interpretations. Hubbell’s neutral theory of ecology is now the standard example of a neutral explanation of statistical models of SADs and other patterns of abundance. There are two popular examples of basic neutral models in ecology: Caswell’s and Hubbell’s. Caswell drew on

²⁴ Some version of this argument is the strongest reason why some ecologists reject the whole problem of inferring process from pattern. They think that there is just not enough information contained in SADs to avoid underdetermination. One reply to this is to note that while many discussions and analyses focus on SADs, niche and neutral theories make many predictions about many patterns. This alone does not avoid underdetermination problems, but it means that they will be *less* underdetermined. See McGill on the importance of multiple pattern testing for theories (McGill 2003, McGill et al. 2007).

work that had been imported to ecology from population genetics to construct his neutral models.²⁵ Hubbell's early work on the community drift model did not cite this work, but his work on the neutral theory did.²⁶ I present the neutral models in this order.

The term “neutral” in biology is used to mean several different, related things. Both Caswell (Caswell 1976) and Hubbell (Hubbell 2001) describe their work as “neutral”. Chapter 2 will take up properly the use of a “neutral” model in Hubbell. The following characterization of neutral models does not fit all “neutral models” in biology. But it does fit Caswell and Hubbell (and Kimura).

In contrast to assuming that species abundance tracks niche space and where resource utilization is assumed to be based on fitness differences between species, neutral interpretations exclude fitness differences and instead track demographic and dispersal processes.²⁷ The abundance of a species in a community can be tracked over time by considering processes which increase, decrease, and do not change the abundance of a species. Additionally, in a *neutral* model these processes of changing abundance treat all

²⁵ See (Caswell 1976, 331) for the lineage he cites. Interestingly, he notes earlier work in ecology that predated population genetics, and does so in a way that suggests he was unaware of this literature before publishing the paper.

²⁶ For this legacy, see chapter 2, section 3.

²⁷ Strictly speaking, a neutral model need not include demographic or dispersal processes. But in practice they always do, because neutrality alone would be really boring. See (Fox 2012) for this “zombie idea”.

individuals, regardless of species, equivalently.²⁸

Caswell built upon work done in and imported from population genetics theory. Both population biology and community ecology are concerned with the diversity of individuals in a group and with the processes producing that diversity. In community ecology the group is a community of multiple populations of different species; in population biology the group is the genome of multiple populations of different alleles. The frequency of alleles in at a locus is analogous to the frequency of species in a community and each is characterized by their relative commonness and rarity. According to the neutral theory of molecular evolution developed by Kimura, Ohta, Crow and others, the observed allele frequencies are the result of the processes of random drift, migration, and mutation. These are analogs of ecological drift (random death and birth), immigration, and speciation.²⁹

Caswell presented three neutral models targeted at SADs and SARs. All three the models describe the probability that the current frequency of a species would increase, decrease, or stay the same over time. These transitions were interpreted in terms of the ecological processes of death, birth, immigration, and speciation. In all three models he assumes all individual's chances of dying, giving birth, immigrating, or speciating are the

²⁸ Caswell's neutral models make this assumption. About neutral models he says, "In such a model the entire set of forces competing for a place in the explanation of the pattern are eliminated. The resulting model is neutral with respect to those factors; naturally, it may or may not be neutral with respect to other factors. The neutral model is used to generate a predicted pattern, which is tested against observation." (Caswell 1976, 330) Assuming neutrality means assuming the absence of the competitive differences. But Caswell's use of "neutral" is focused on the flip side of the coin. I find that he uses "neutral" to mean what I call the use of a neutral model as a "null model" - its ability to test a competitive model. Caswell emphasizes this use, while Hubbell emphasizes the assumption itself.

²⁹ See (Leigh 2007) for the best historical comparison of the neutral theory of molecular evolution and the neutral theory of ecology.

same across species. This is explicitly done to neutralize the effects of selection and competitive niche exclusion.

Model I describes a community that is drifting, whose birth rate is equal to the death rate, and to which individuals of novel species can immigrate. This is the relaxed version of assuming that the community size remains constant, where on average deaths are met with births. Immigration is the source of diversity as birth and death alone can never increase decrease diversity. Model II was the same except that population size remains constant and every death is met immediately by a birth or an immigrant. Finally model III also imposed constant community size and included a process analogous to mutation. Here, every turn a random sample with replacement is taken of community at the previous turn and there is some chance that any of these individuals sampled will “mutate” into a novel species. This chance is determined by the binomial sampling probability and a mutation rate (Caswell 1976, 332).³⁰

Caswell reported that the expectations of neutral models are robust across these differences in assumptions. He asserted that each of the three models, while having slightly different dynamics, produces the same: number of species, approximate SAD, and sampling properties of the SAD.³¹ He also surveys similar models proposed by

³⁰ With respect to Hubbell’s origin story, one interesting observation here is that in models I and II, what is interpreted as immigration in the ecological models was interpreted as mutation in the evolutionary models, while model III retained the interpretation in terms of “mutation” - which is also interpretable as speciation! The only time Caswell mentioned “speciation”, he said “speciation or invasion”, and so I think that he is not concerned with the difference and that for him the model allows either interpretation.

³¹ (Hubbell 2001, 51) questioned whether all three models really produced the same distribution. Hubbell expected a different distribution resulted from holding the community size constant.

others, and explains that all the models produce Fisher's logseries distribution³².

Caswell framed his discussion of the mathematics of the models in terms of MacArthur and Wilson's Theory of Island Biogeography, because of the emphasis on immigration rates. The theory of island biogeography offered a completely different understanding of the processes and causal factors responsible for the SADs than did the tradition started with MacArthur's broken-stick model with its interpretation in terms of niche exclusion. The difference between the project started by MacArthur's niche apportionment models³³ on the "niche" or "competition" side, and the Theory of Island Biogeography on the "dispersal" or "neutral" side, runs all through community ecology. It will feature heavily in chapter two's second narrative focused on why Hubbell calls the theory "neutral".

Hubbell presented the *community drift model* of SADs in (Hubbell 1979, Hubbell and Foster 1986). This model will become a parent of the neutral theory of ecology in (Hubbell 2001). Relatively little space in either of the two early papers is devoted to the community drift model. Hubbell quickly mentions that Caswell's work on the "neutral" models did not produce the lognormal distributions that best fit sampled data. He then proposes that his model can be thought of as a "dynamic version" of MacArthur's broken stick model (Hubbell 1979, 1306), which brings it in line with the perspective taken by MacArthur and Wilson's work on the theory of island biogeography's and its focus on extinction and immigration (Hubbell and Foster 1986, 323).

³² Caswell imports results proved in population genetics including (Watterson 1974).

³³ And also the project started by MacArthur's work on a graphical-mathematical theory of competition (MacArthur 1972). This is discussed in chapter 5.

Hubbell used the same community drift model in both the 1979 and 1986 paper. Begin with a random composition of individuals in an area with fixed community size K . Let a disturbance come and kill D individuals, where the probability that a species decreases in size is proportional to the previous relative species abundance of that species.³⁴ Then, to keep the community size constant, replace the dead with D new individuals. The probability that an individual of a given species will establish itself is proportional to the abundance of that species in the community *after* the disturbance. Then repeat this over and over again for a given number of turns.

Hubbell determined the properties of this model varying K , D , and the number of turns, doing Monte Carlo computer simulations, and averaging the results of many trials. For a given K and D , in the short term (25 turns), the lognormal-like distributions were obtained. In the intermediate-term (250 turns), geometric-like distributions were obtained. In between these time scales, logseries distributions were obtained. Finally, in the long-term, the diversity of the community completely vanishes and there can be only one species in the community, since species can go extinct but not originate. In the short term through computer simulation. Most originally, the model showed a tail-off phenomenon, where the least common species are rarer than expected by the logseries or lognormal distribution.

Hubbell then went on to augment the basic model with immigration to balance out extinction and so maintain diversity like the theory of island biogeography. In the paper he only presented this augmented model verbally, without equations, though the results

³⁴ This will of course be retained as the “neutrality assumption”, though Hubbell never identifies it explicitly as an assumption and certainly never calls it “neutrality” until the book. See chapter 2 for a discussion.

that he claimed for the augmented model implied that he did simulations. By varying the values of the immigration rate and the community size, this model could now produce time-independent, equilibrium patterns from lognormal to geometric, and these distributions were stable in time in the sense that, after some number of turns of the model, the expectation for the SAD is approximately (because the model is stochastic) fixed for all later times.

Hubbell presented a modified version of the community drift model in (Wright and Hubbell 1983). Rather than directed at modeling SADs directly, it targeted the SLOSS (single-large or several-small) problem of natural reserve design. Wright and Hubbell argue that the answer depends crucially on the presence and strength of immigration to the reserve. This model is *density-dependent* community drift model. In the basic density-independent community drift model, the probability that an individual of species s will replace a death is proportional only to the relative abundance of species s . The density-dependent community drift model singles out one *focal species* to have different birth and immigration rates than the rest of the species. This is an augmentation of the community drift model which retains neutrality. The birth rate and immigration rate of the focal species depend upon its current population size. If the focal species is near its carrying capacity, then its chances of birth and immigration are lower than if its population is small. Wright and Hubbell do not say which SAD is expected from the density-dependent community drift model.³⁵

³⁵ This case is most interesting because it hints at the research programme that Hubbell will found with the book, and which I will characterize in chapter 5. (Volkov et al. 2005) adds density-dependence to the neutral theory, but here it is symmetric and treats all individuals the same. Therefore, it obeys neutrality. But arguably it is no longer a purely demographic and dispersal theory because density-dependent growth may be caused by interspecific competition.

Hubbell does not publish on the community drift model again until (Hubbell 1995). Here he proposes adding speciation as the ultimate source of diversity. In (Hubbell 1997) he presents the basics of the “unified model”, and in the book (Hubbell 2001) presents the basic neutral model described above in section 1 above.³⁶

While Hubbell’s earlier did not cite work done in population genetics, he explicitly incorporated and cited these results when working out the full theory presented in the book. The neutral theory is a metacommunity theory with two scales. The metacommunity can be thought of as the mainland where speciation and extinction are occurring and which sends immigrants to the island. The neutral theory gives one distribution for the metacommunity and another for the local community, and each can be compared with SADs. Hubbell calls both of these the *zero-sum multinomial distribution*. In working out what distribution was expected in the metacommunity, he drew on work done in the analogous problem in population genetics solved by Ewens and Karlin - the same sources Caswell had drawn on (Ewens 1972, Karlin and McGregor 1972, Hubbell 2001, 119, Caswell 1976).

Hubbell shows that the metacommunity equilibrium distribution depends upon one parameter, θ , called the *fundamental biodiversity number*, which is a function of metacommunity size J_M and speciation rate v , and is logseries-like. This makes sense since Caswell’s models only had immigration or speciation and drift going and produced

³⁶ In between (Bell 2000, 2001) were published. Bell’s papers are important, but not for this narrative. Bell’s models do not differ much from Hubbell’s community drift model. Bell’s papers play an important role of chapter 2 in answering why Hubbell called his theory “neutral”.

a logseries distribution.³⁷ Hubbell drew on a result proved by (Watterson 1974) in the analogous problem in population genetics to show the metacommunity distribution asymptotically approaches the logseries as metacommunity size goes to infinity (Hubbell 2001, 125). One lasting use of Fisher’s logseries has been using its parameter α as a phenomenological measure of diversity meaning roughly the number of species with one member. Hubbell’s result that $\alpha \cong \theta$ shows that, assuming truth of the neutral theory, that the number of singleton species is related to the product of the metacommunity size and the speciation rate, giving it an biological interpretation (Hubbell 2001, 165).

Finally, a lognormal-like distribution is obtained for the local community. The local community drifts and receives immigrants according to the relative abundances in the metacommunity and therefore depends upon θ and also local community size J and immigration rate m . The local community ZSM distribution is lognormal-like, except it expects rarer species than the lognormal. Finally, reducing the immigration rate produces a local community ZSM which is geometric-like that has fewer species of less evenness. Hubbell praises the neutral theory ability to unify the previous disparate distributions³⁸ and offer the explanation that they differ in the ecological scale they describe and the connection between the scales.

5.5 Lingering epistemic issues

³⁷ This shows that the zero-sum multinomial distribution was ill-named. Hubbell thought at the time that assuming constant community size was necessary to get ZSM dynamics. But this was shown unnecessary in (Volkov et al. 2003), and was anticipated by the claimed robustness of distributions produced by Caswell’s model’s I and II.

³⁸ This is not the primary meaning of “unify” in “the unified neutral theory”. The primary thing unified by the neutral theory is biodiversity and biogeography, with, for example, SADs and SARs.

The expansion of Hubbell's origin story, considered as a historical claim, that I have told in section 5 shows that Hubbell sees the development of the neutral theory as primarily a response to the project of constructing a model which produces a statistical distribution that fits SADs and other patterns well. Previous models had various limitations, especially either not producing accurate statistical distributions, not having a biological interpretation, and not having measureable parameter values. Adding speciation to the community drift model based on drift and immigration helped to make progress on this problem. I have shown how the project has at least four levels - data, representation of data, statistical distribution fit to representation of data, and interpretation of statistical distribution fit to representation of data - and also how each transition poses significant epistemic difficulties for inferring processes from patterns. The main epistemic difficulty is the underdetermination of theory by data problem - more than one theory is capable of making the data expected. In the case of community ecology, more than one interpretation can always be given for any statistical distribution, more than one statistical distribution can sometimes be fit equally well to a given representation of data, and the data can always be represented in more than one way. This means that while the neutral theory might make the abundance data likely and so we can sometimes infer the data from the theory, we cannot uniquely infer the theory from the data.

But the situation is even more complicated than I have explained, and so before closing this section, several lingering epistemic issues need to be mentioned before moving on.

First, I have not discussed methods for evaluating the fit between a statistical distribution and a SAD. Measuring the fit between a statistical distribution such as the lognormal distribution or the broken-stick and a SAD is an additional problem. There are two aspects of this problem. Part A is how to measure the distance between two functions. Part B is how to determine the parameter values that best fit one function to another. If you have a curve that you want to fit using a particular function, then some sets of values of variables (parameter values) will make the difference between your function and the curve you aim to fit smaller than other sets. Of special interest is the, possibly non-unique, set of parameter values which minimize the difference between a function and a curve. There are multiple ways of doing both part A and B. In practice, two ways of approaching part B are linear regression and maximum likelihood. These procedures determine the what values of the parameters of the model would give the best model-fit to a given SAD. This is much more complicated for stochastic models than deterministic models. Obviously, showing that a model is capable of good model-fits is significantly easier than showing that a model fits the data well given actual independently measured parameter values.

Second, I have not discussed *testing* statistical distributions or interpretations of statistical distributions. While the most common analyses of statistical distributions such as Hubbell's ZSM are based on best model-fit parameter values arrived at by regression or maximum likelihood methods, it is not agreed upon that this is useful for *testing ecological hypotheses*. Determining best model-fit is all that can be done when using statistical distributions phenomenologically, without interpretation, because the

parameter values have no interpretation, and are not used to make inferences to process. Statistical distributions can be compared for their ability to fit some data though, and in this way tested towards the goal of finding the best best-model-fitting distribution.

Methods used here for comparing statistical distributions are maximum likelihood ratio tests and the Akaike information criteria. These are also potentially useful for making predictions for how a SAD will vary with space and time. But if processes are to be inferred from patterns, these methods are inadequate. The ideal way to test hypotheses requires independently measuring the parameter values and comparing the tuned distribution to a SAD. But this is difficult for two reasons. Reason A is that even *measuring* immigration rate is difficult, let alone speciation rate. Additionally, when an immigration rate is measured, as is often done using seed traps or counting birds, the model must be calibrated properly to accept some known transformation of these numbers. Reason B is it is not obvious how to *independently* know how parameters change in space and time. This limits extrapolating from one measurement of parameters values at a place and time to other places and times. Right now, one method used for estimating parameter values is though using linear regression and maximum likelihood methods. Obviously applying these parameter values to the same data model as they were estimated from shows only internal consistency. But if they could be applied to new data models, this would be some independence. But this requires knowing how immigration and speciation rates change in time and space.

Testing hypotheses also raises other possibilities for gaining evidence independent of model-fit for a statistical distribution or model, such as whether the assumptions of the model are met. This is rarely done because, I assume, it is much more

energy intensive and difficult than using formal methods such as the Akaike information criteria.

Third, I have not discussed the dependency of model-fit with the neutral model and the spatial *scale* at which the abundance data is collected. Abundance data is collected for scales starting at the very small, e.g. 1m² plots, on to the continental scale. Most whole communities reside in between these extremes. The shape of a SAD has been shown to depend on scale, and therefore the model-fit and the outcomes of hypothesis testing and other analyses using model-fit will also depend upon scale.

Fourth, I have not discussed whether neutral interpretations of statistical distributions of SADs are *biological* interpretations, or *statistical* interpretations, or *something else*. Opinions are divided when it comes to categorizing neutral models. For example, Magurran categorizes the logseries and lognormal distributions as *statistical distributions*, all the niche apportionment models as *biological (or theoretical) models*, and Caswell and Hubbell's neutral models as *other approaches* (Magurran 2004, chapter 2). Caswell thought of his neutral models as *neutral* because they were null with respect to species interactions posited by niche apportionment models. This is more like the absence of a biological interpretation, and so a statistical interpretation. Hubbell uses *neutral* to mean identical individual ecological rates, which is more a biological interpretation. In short, neutral models are neatly neither only biological nor statistical models, and are capable of both interpretations (Gotelli and McGill 2006). The important difference is not so much in the content of the models, but in the way that they are used. I

discuss the way this affects the way neutral models are reasoned with in chapter 2, 3, and 4.

6. Conclusion

Section 5 filled out a part of where the neutral theory of the book came from - the question that Hubbell's origin story is meant to address. Hubbell says that the development of the community drift model between 1986 and 1995 was spurred in 1995 by the question from the student in his class. I have shown in this chapter that it is highly misleading to suggest that Hubbell, prompted by the fateful question, unfolded the neutral theory of ecology from the Theory of Island Biogeography.

Hubbell's origin story does emphasize what Hubbell sees as the distinguishing characteristic of the unified neutral theory, namely its abilities to produce both logseries-like and lognormal-like distributions, as well as predictions about biodiversity and biogeography. This resulted from introducing a metacommunity in which speciation supplies diversity via immigration to a local community under drift. This can retrospectively and conceptually be understood through augmenting the Theory of Island Biogeography, but it only actually happened given Hubbell's earlier work on the community drift model, which did not build upon the Theory of Island Biogeography. Hubbell's origin story rewrites the ancestry of the neutral theory. Moreover, if a conceptual origin story is sought, we might just as well emphasize the role of the neutral theory of molecular evolution as a source of the neutral theory of ecology.

I critique Hubbell's origin story on historical grounds, not literary grounds.³⁹ Even granting that the question *did* spur Hubbell back to further developing the model he had begun 14 years before, the stage had been set for such a spark through his long involvement with the modeling project this narrative foregrounds. This is the history which is glossed over in the eureka moment his student's question induced.

I began this chapter by explaining the received way of viewing the controversy surrounding the neutral theory. Ecologists fixated upon the truth or falsity of neutrality as a formal assumption. The worry was that a false foundational assumption should undermine the ability of a theory to be explanatory or useful. But the existence of the neutral theory shows that a false assumption can be used to construct a model which is empirically adequate and so potentially explanatory and useful. The proper way to address this tension seems to be to get analyzing the way in which false assumptions are actually useful and can be explanatory. Since there are different way of being false, some being worse for certain purposes than others (Wimsatt 1987), work should be done to see the sense in which neutrality is false.

Hubbell's origin story, and the expanded narrative I presented above, buttresses this characterization by showing that the reluctance of ecologists towards neutrality stems from the general commitment to the importance of interspecific competition and competitive exclusion as the dominant cause of patterns of abundance and diversity. The closest thing that ecology has to a paradigm or dominant research programme is centered

³⁹ Historians of science are divided on whether historically inaccurate stories such as Newton's apple are *good stories*. The split is between those to see a value to such origin

on competition theory. This enriches the epistemic situation by showing that the issue is not only internal to the neutral theory, but also concerns its relationship to competition theory. The focus on constructing models of abundance patterns such as SADs on the basis of ecological hypotheses further showed that this problem is massively underdetermined, many models producing empirically adequate predictions. The conflict between competition theory and neutral theory is an instance of this general underdetermination problem, and their disagreement is purely at the level of hypotheses and theory.

This way of seeing the epistemic situation focuses only on the model-fit of the neutral theory. While I explained how the community drift model developed into the neutral theory, the underdetermination problem arrived at only looks synchronically at the neutral theory. No attention is paid to other potential sources of support for neutral theory. Only accurately describing patterns such as SADs using mathematics is discussed, and this portrays the neutral theory very abstractly, with little connection to actual communities.

In the next chapter, I investigate two more historical questions about the origins of the neutral theory. The narratives I give to answer these questions enrich and shift our understanding of the epistemic situation. The later three chapters investigate epistemic questions that arise from, and are informed by, this enriched history of the neutral theory.

myths besides historical accuracy and those who think that the image of scientific innovation such stories teach are harmful to understanding science.

Chapter 2.

Retelling the Origins

1. New narratives for new philosophies

Contrary to what you have been led to believe, both by Hubbell and myself in the previous chapter, the neutral theory of ecology did not develop primarily as a model of relative species abundance distributions. And neutral theory is not controversial only because, or even primarily because, it assumes neutrality and hypothesizes demographic and dispersal processes, instead of interspecific competition. In this chapter, I inquire into two new historical questions about the origins of the neutral theory. In section 2 I ask, What was Hubbell working on before he wrote his book on the neutral theory? In section 3 I ask, Why did Hubbell call his theory “neutral”?

Answering these questions enriches our understanding of where the neutral theory came from and why it is controversial. In section 2, I show that neutral theory is controversial because it is tied to a debate larger than the competition vs neutral theory debate, namely debate over whether communities are in taxonomic equilibrium, and so maintain the same group of species over time, or nonequilibrium. In section 3, I show that neutral theory is also controversial because it is inextricably tied to the methodology based on using neutral models as ‘null hypotheses’, which has a messy and contentious history across biology going back to the 1970s. In addition to enriching the history, these chapters reframe the epistemic questions I opened chapter 1 with. They also provide new resources for analyzing and answering them.

2. What was Hubbell working on before he wrote the book? Tropical Forest ecology

Hubbell presented the community drift model in the context of evaluating the dominant community ecology problem of the day, whether communities are in (taxonomic) equilibrium or nonequilibrium. I explain the difference between (taxonomic) equilibrium and nonequilibrium communities at the beginning of section 2.1. What Hubbell did was formalize a prior verbal nonequilibrium hypothesis and use it to explain how it was possible that a particular tropical dry forest census was in nonequilibrium, a conclusion he independently supported with direct evidence from an examination of the system.

In this section I show how the community drift hypothesis was a tool in this much larger equilibrium controversy. A significant amount of the passion for or against the neutral theory today stems, not from its particular ability to accurately fit SADs, but from the support it gives to the nonequilibrium nature of especially tropical rain forest and coral reef communities. The equilibrium problem was the motivating problem for the community drift model, the immediate ancestor of the neutral theory, and the problem that connected Hubbell's theoretical work on neutral theory and observational work on tropical forest dynamics.

I begin by showing that the context of Hubbell's first two papers featuring the community drift model is the equilibrium vs nonequilibrium problem. In section 2.2 I

give historical background to this question by explaining the views of Jared Diamond and David Lack on island biogeography. In section 2.3 I give a rereading of Hubbell's early papers with this problem foregrounded. In section 2.4 I investigate the coincidence of Hubbell's work founding the Barro Colorado Island Forest Dynamics Plot project and his work on the community drift model.

2.1 The (Taxonomic) Equilibrium Problem

Hubbell's 1979 paper evaluated the Janzen-Connell hypothesis for a particular census of a tropical dry forest in Panama. The Janzen-Connell hypothesis is an explanation for "the widely held generalization about tropical tree species ... that most occur at very low adult densities and are of relatively uniform dispersion, such that adult species are thinly and evenly distributed in space" (Hubbell 1979, 1299) In 1970, Daniel Janzen and Joseph Connell independently proposed very similar explanations of the thin and even distribution of trees in a tropical forest. Janzen proposed that the seeds that trees drop near their base will be eaten by animals who like to eat those seeds. Seeds which are further away from the parent and not clustered will have better chances of not being eaten and so germinating. Hubbell explained it as, "Janzen and Connell both argued that such host-specific attack by herbivores would reduce the local density of any given species, open up habitat to invasion by additional species, and thereby maintain high species diversity" (Hubbell 1979, 1299).

In addition to evaluating the Janzen-Connell hypothesis for one tropical forest, Hubbell was concerned with the more general question of whether communities are in equilibrium or nonequilibrium. The Janzen-Connell hypothesis was one particular

equilibrium hypothesis for the widely accepted empirical generalization that the distribution of individuals in a tropical forest is thinly and evenly distributed with respect to species. Connell elsewhere explained the distinction between equilibrium and nonequilibrium hypotheses of observed abundance and diversities.

“Equilibrium of species composition is usually defined as follows: (i) if perturbed away from the existing state (equilibrium point or stable limit cycle), the species composition would return to it; (ii) without further perturbations, it persists in the existing state. A perturbation is usually regarded as a marked change; death and replacement of single trees or coral colonies would not qualify.”

(Connell 1978, n1 on 1309)

This concept of taxonomic equilibrium is distinct from the concept of time-independent equilibrium. The distinction between *taxonomic* equilibrium and nonequilibrium concerns how a community's species composition responds to disturbance. SADs represent some of the information about a community's structure, but they include neither the identity of different species nor the relative abundances between particular species. Therefore, a set of some temporal series of SADs is, by itself, compatible with both equilibrium and nonequilibrium explanations, because the equilibrium status of the community depends upon whether specific species continue to exist over time. The distinction between *time-independent* equilibrium and nonequilibrium concerns whether the solutions of a theory's equations are time-independent. The theory of island biogeography arrives at the (time-independent) equilibrium number of species on an island by solving for the species richness when the balance between immigration rate and extinction rate does not change with time. There is no necessity in doing this, but it is a common assumption made for both mathematical and biological reasons. Mathematically, time-independent solutions are much simpler.

Biologically, using only equilibrium solutions means assuming that the number of species in the community is no longer changing with time. Approach to equilibrium solutions could also be worked out. The informal characterization of the neutral theory in terms of the checkerboard and the rules to change the board over time show how the time-evolution works. This is what Hubbell did with the drift-only community drift model via computer simulations. By varying the number of turns, he showed how the SAD changed from lognormal to geometric, and so these were not time-independent equilibrium solutions. The full neutral theory of ecology, as presented in the 2001 book, is a time-independent equilibrium theory however. Both the community drift model and the neutral theory are regarded as *taxonomic nonequilibrium* theories because they make no assumptions about species composition and assume that species are going extinct and novel species are immigrating.⁴⁰ Unless otherwise noted, in what follows “equilibrium” refers to taxonomic equilibrium.

Connell presented 3 particular nonequilibrium and 3 particular equilibrium hypotheses for how diversity is maintained and then evaluated the different hypotheses for different communities (Connell 1978). One hypothesis stands out - the nonequilibrium *equal chance hypothesis*:

“Species are approximately equal in ability to colonize, exclude invaders, and resist environmental vicissitudes. Local diversity depends only on the number of species available in the geographical area and the local population density” (Connell 1978, 1303).

Connell’s paper goes on to examine whether any communities exist that meet the needed

⁴⁰ There is actually a third distinction used by ecologists: *quasi-equilibrium*. A community is in quasi-equilibrium if it maintains a stable SAD over time, and has slow species turnover. The neutral theory is a quasi-equilibrium theory in this sense. (Hubbell 2001, 70-71).

conditions to support this hypothesis. He noted that support had been given for this view by several ecologists, one studying coral reef fish and one studying rain forest trees, and called for further tests to be done. He said though that, “it seems unlikely that either rain forest trees or corals conform to the equal chance hypothesis.” (Connell 1978, 1306)

Hubbell’s community drift model is an instance of the equal chance hypothesis, but Hubbell extends neutrality from the species level to the individual level. Hubbell was clear about this, saying that the community drift model, “... in its simplest form as presented here corresponds to the “equal chance hypothesis” discussed by Connell, provided that per capita chances of reproduction or death are made the same for all species” (Hubbell 1979). Connell’s explanation shows that he is also speaking of a demographic and dispersal-based hypothesis.

Hubbell constructed a formal model of this verbal hypothesis. While Hubbell did not make the connection explicit, assuming equal chances is just assuming *neutrality*. Remember, neutrality means that all individuals, regardless of species, are ecologically equivalent. The community drift model states that the SAD at a time just depends upon the relative abundances of species in the local and surrounding areas. Putting neutrality and demographic and dispersal processes together, the SAD at a time *only* depends upon the relative abundances of species in the local and surrounding areas. In this sense, it hypothesizes that abundance patterns result from history and chance.

Hubbell claims he was not aware of Caswell’s work on neutral models when doing this work and was not thinking of the work in population genetics.⁴¹ Hubbell’s first

⁴¹ Personal communication, 2015. Hubbell says that he became aware of it when looking to publish the paper.

paper on the community drift model needs to be reevaluated in light of the equilibrium problem. But before that, the larger context of the equilibrium problem at the time is needed.

2.2 The Larger Context of the Equilibrium Problem: Jared Diamond and David Lack

In 1975, Jared Diamond published the massive “Assembly of species communities” (Diamond 1975), wherein he argued that communities are structured by *assembly rules*. This paper caused tidal waves through community ecology. He opens the paper with his position stated very clearly,

“This chapter explores the origins of difference in community structure, such as those between islands of the same archipelago, between different localities on the same island, between adjacent habitats, and between different biogeographical regions. The working hypothesis is that, through diffuse competition the component species of a community are selected, and coadjusted in their niches and abundances, so as to fit with each other and to resist invaders. Observations are derived from bird communities of New Guinea and its satellite islands, of which some are at, some above, and some below equilibrium species number (S).” (Diamond 1975, 342)

Diamond aimed to explain the origins of differences found in community structure, and he proposed to do so with a hypothesis inducted from data concerning bird communities. His data allegedly showed the various species composing a community do not just happen to be found together in the same place and at the same time, but rather form a cohesive and interconnected whole. For this reason, Diamond argued that the birds of Melanesia were in taxonomic equilibrium. The details of his reasoning, and the particular assembly rules that Diamond proposed, need not concern us now, but they will show up

again in section 3.3 below. The important point is that Diamond's work excites and directs attention towards the issue of equilibrium vs nonequilibrium communities.

Even though MacArthur and Wilson's theory of island biogeography explains the number of species on an island (S) as the result of the balance between immigration rate and extinction rate, it is usually interpreted as a nonequilibrium hypothesis that predicts continual species turnover in the sense that residents will go extinct and be replaced by new migrants.⁴² There is no requirement in the theory that, for example, already established species are adapted to the island's conditions and so can resist new migrants and in this way maintain a constant species number.

David Lack proposed just such a view in his 1976 *Island biology, Illustrated by the land birds of Jamaica*. Lack's work is relevant in three ways to us here. First, Lack countered the theory of island biogeography's nonequilibrium explanation on its own grounds. Lack agreed with MacArthur and Wilson that a phenomenon to be explained included the observation that "Oceanic islands never connected to the land have far fewer resident species of land birds than the mainland (per area), and fewer the smaller and more remote that they are." (Lack 1976) But Lack additionally included a further observation, not significant in the basic version of the theory of island biogeography, that "Various of the island species also have broader ecological niches than their mainland counterparts, and many are endemic [found exclusively on that island]."

Lack explained these observations as follows. The number of species on an island

⁴² This question is complicated by the fact that there were two authors and both authors continued to work on the ideas that began here. MacArthur is said to have ended up with the position that continual species turnover is not required and even not expected on islands even where the immigration rate is a difference maker. This comes out in the preface to (Lack 1976). TIB is at least a-equilibrium, agnostic on the issue. Hubbell interprets the TIB as a nonequilibrium theory and as a *neutral* theory at the species level.

depends upon the *ecological poverty* of the island. Less poverty supports more species. Islands are ecologically poorer than mainlands, more distant islands poorer than closer islands, and smaller islands poorer than larger islands. The simplest reason for this is that smaller islands tend to have fewer and smaller habitat types than larger ones and species need a minimum area to be able to live. This leads species on an island to expand their ecological niche and become more generalist than their ancestors were - 'fewer generalists exclude more specialists'. Lack offered no hypothesis about *why* more remote islands are ecologically poorer.⁴³

In studying Jamaica and the Caribbean islands in detail, Lack found that birds fly to an island much more frequently than the theory of island biogeography requires. Lack saw birds constantly streaming back and forth between the islands and the mainland. He hypothesized that most could not get a foothold because the spaces were already almost full. What the theory of island biogeography explained with the rarity of immigration/dispersal, Lack explained with ecological poverty.

Both Diamond's and Lack's proposed a form of *ecological determinism*. On this view diversity patterns are determined largely by the particular interactions, mediated by natural selection via competition and adaptation, of different species and the abiotic conditions. Contingent historical events such as the order of colonization of an island play a small role. This was in contrast with the view of the theory of island biogeography where diversity is a product of chance and historical processes, and communities are random assemblages of species which are not strongly and consistently interacting with

⁴³ I'll here just note that Lack wrote the book after he was diagnosed with cancer and in ill health and died before it was published. In a sad coincidence, MacArthur had also

each other and whose diversities are not predictable from knowledge of the biotic and abiotic relationships.

The second relevant feature of Lack's work is that Lack and MacArthur and Wilson are distinguished by whether they take a *close* or *distant view* of the communities they targeted. Lack used Williams' (Williams 1969) distinction of close vs distant views on ecological problems.

“The ‘distant’ view depends upon a priori ideas, mathematical analyses and selected examples, and in the hands of masters has led to outstanding advances. But it has two weaknesses in less hands, first a tendency to concentrate on numbers of species rather than the particular species involved, and secondly, because nature is so various, one can with diligence find examples to illustrate almost any new idea. The ‘close’ view, in contrast, depends on the intensive study of a particular situation in all its aspects, from which conclusions are later drawn.” (Lack 1976, 7)

Lack took a close view and drew out knowledge of island biogeography from the intensive study of birds on Jamaica. MacArthur and Wilson took a distant view of island biogeography and drew out knowledge of island biogeography from the study of very general ecological patterns and processes. Lack conceded that he might have just preferred to work in the close method, and that he reread MacArthur and Wilson's book more times than any other. But he clearly took some pride in taking the close view.

The third relevant feature of Lack's work is that a strong correlation emerges between what side of the equilibrium vs nonequilibrium debate you come down on and what type of organisms the community you study is made up of.⁴⁴ In particular,

been diagnosed recently with cancer, quickly wrote his *Geographical Ecology* (MacArthur 1972), his final statement on all matters, and died shortly before Lack.

⁴⁴ This point was made by (Orians 1962) and discussed by (Cooper 1993). Simberloff also makes this connection somewhere, but I cannot find the reference.

equilibrium views tend to come with studies of bird communities and nonequilibrium views tend to come with studies of tropical forests and coral reefs. One biological reason for this is that birds can move while trees and corals cannot. But the connection is also historical and disciplinary. We have seen this correlation with Diamond, Lack, and Connell, and it will be true of Hubbell. MacArthur and Wilson are the outlier here because MacArthur studied warblers and together they used bird data.

We have now three distinctions: equilibrium vs nonequilibrium community, close view vs distant view, and bird communities vs tropical forest communities. We almost have all the pieces needed to reread Hubbell's early papers and understand their context and the role played by the community drift model.

A final piece is that Connell's 1978 paper, which I used to introduce the equilibrium problem in section 2.1, accepts that all of the different equilibrium and nonequilibrium hypotheses he considers are mutually compatible and are each certainly happening sometimes. Hence the ecological problem was properly to figure out the relative significance of each hypothesis in different communities, not to test mutually exclusive hypotheses and eliminate all but one. This relative significance problem required much better study and data, in particular large scale data, of communities.

2.3 Rereading Hubbell's early papers in light of the Equilibrium Problem

Hubbell's 1979 paper evaluated the empirical generalization that the Janzen-Connell hypothesis was proposed to explain - trees are thinly and evenly distributed with respect to species in a tropical forest. It also evaluated explanations of tree rarity and the equilibrium status of a tropical dry forest in Costa Rica. The study site was a 13.44 ha

plot, within which every woody plant with a diameter at breast height (dbh) greater than 2cm was (1) identified by species, (2) measured for dbh, and (3) mapped spatially within a 1m x 1m grid. I will not present Hubbell's detailed analysis of all the data for this plot because it is more interesting to do it for the data used in the later papers. Against the Janzen-Connell hypothesis (one equilibrium hypothesis) Hubbell found that tree species are not uniformly distributed but are somewhat clumped (this depends upon species, scale, and how adults and juveniles are handled). He further gave direct evidence for the forest might being in nonequilibrium. His clearest message was that evaluating these claims required much better data, in particular *dynamic* data of how the forest changes through time and plots of longer temporal scale need to be begun.

The use of the community drift model was not used as a model of SADs to infer the underlying mechanisms. This is clear from the three lessons Hubbell drew from the use of the community drift model. Remember, the model here is a nonequilibrium model that tracks individuals under ecological drift and a free immigration parameter.

“First, we may expect to observe substantial differentiation of the relative abundance of species in natural communities as a result of purely random-walk processes—a kind of “community drift” phenomenon.” (Hubbell 1979, 1307)

By varying the time scale of the model and the strength of the immigration rate, quite differently shaped distributions are predicted by the model. This is a kind of proof of principle of the versatility of the model.

“Second, we cannot necessarily conclude that, just because a species is of rank-1 importance in a community, its current success is due to competitive dominance or “niche pre-emption”, stemming from some superior adaptation to the local environment. Because such a simple model generates the basic patterns of relative species abundance in natural plant communities, it would perhaps be

preferable to use departures from the lognormal or geometric distributions as evidence for competitive dominance.” (Hubbell 1979, 1307)

The ability of the model to qualitatively reproduce observed SADs undermines the view that observed SADs are necessarily from a community in taxonomic equilibrium where coadaptive mechanisms are dominating. Further, rather than using the ability to fit SADs with lognormal or geometric distributions as evidence for competition, as Diamond did with his assembly rules, Hubbell suggests that it is *departures* from these distributions that give evidence and maybe even allow quantitative measurements of the strength of coadaptive mechanisms.

“Finally, whether or not forest communities are at equilibrium, our understanding of community organization would profit from more study of processes of disturbance, immigration, and local extinction, in conjunction with the more traditional studies of the biotic interactions of species (such as competition, niche differentiation, and (seed predation).”(Hubbell 1979, 1307)

Finally, the model is again a proof of principle of a quantitative model that does not follow the lead of Diamond and Lack, but rather of the theory of island biogeography.

The original context of the community drift model comes through even more clearly with a rereading of Hubbell and Foster’s 1986 paper. Its stated purpose concerns “the structure and dynamics of tropical rain forest tree communities, in particular, why chance and history might be expected to play an especially important role in structuring these communities.” (Hubbell and Foster 1986, 314) Against ecologists who hold communities to be at or near equilibrium and who use competitive niche exclusion and explain diversity, they argue that tropical tree communities are not near equilibrium, and that events such as wind blowing down trees, seeds just happening to land in the right

place at the right time, and the past relative abundances of species are most causally significant overall.

The study site of this paper is a 50 ha plot on Barro Colorado Island (BCI). Hubbell and Foster led the group that identified and mapped all free standing woody plants over 1cm dbh. From a single census of this plot, they make the following argument: in order for Diamond or Lack's view of species abundances to hold, the forest composition should be the result of highly integrated group of species, each specialists at exploiting a particular niche, all under frequency-dependent selection, and maintaining a stable taxonomic equilibrium. In order for a community to satisfy this view, individuals of different species need to spend a lot of evolutionary time around each other in order to coevolve and specialize. But the BCI data shows that individuals of most tree species are surrounded by many different kinds of trees, and two individuals of the same species have mostly different species as neighbors. Because the neighborhoods trees live in are highly uncertain, they are unlikely to promote the coevolution required for an equilibrium view. The frequency of particular pairs of species being near each other is too low. This is compounded by the different life history stages of trees. In a community like this, selection will not point in a single direction for a species, but different directions for different individuals of the same species. This further explains why many pollinators can pollinate the same species of tree and why one tree species can be pollinated by many species.

This community structure will lead to convergent generalization, not divergent specialization. For example, adult trees shading saplings growing under them will select for shade-tolerance in all species, not for specialists in shade-tolerance differing by

degrees based on species. The same idea holds for soil nutrients. Hubbell and Foster note that tropical tree species have many mutualistic mycorrhizal fungal symbionts in common in the soil that help trees take up soil nutrients such as phosphorus, more easily.

They do find that there are clear differences amongst a few guilds distinguished by their life history preferences. A guild is a group of species that utilize the same resources and live in a similar way. One important limiting resource in the forest is light, and some guilds are distinguished predominantly by the type of light conditions they are found in. The gap pioneer guild first fill the gaps left by falling trees. They are the most heliophilic and also the most transitory. These are replaced by the late secondary heliophiles guild, which composes 25% of the species filling in the canopy, but which is quite rare overall on BCI. Finally, there are several guilds which are shade-tolerant. These are slower growing, but tend to live a long time, some of whom slowly take a place in the canopy.

Most of the trees in the tropical forest do not have a place in the canopy. The most common species are found across varied conditions. Overall Hubbell and Foster find about 12 guilds. They argue that the sophisticated niche theorist would expect each species or guild to specialize in one set of conditions and expect to find a continuum of adaptive strategies. But because 70% of species showed no specialization in regeneration requirements, they conclude that most species are “for all practical purposes, functionally identical (Hubbell and Foster 1986, 318).

To summarize: they argue that interspecific competition does not invariably lead to specialization, but often leads to generalization and similarity among guild members, particularly if selection is diffuse and in many directions. Species-rich communities with

high unpredictability in the biotic environment may cause diffuse coevolution of guilds within which members are functionally equivalent. In these conditions, chance and history understood as initial local and biogeographical distributions have a much stronger effect on SAD than interspecific competition.

Hubbell and Foster hypothesize that many tropical tree species coexist *because* they are functionally equivalent generalists, not *in spite of* this, as the equilibrium schools think. They give three lines of evidence for this hypothesis. The first of these is Hubbell's community drift model. The complete use of the community drift model is to establish that, "ecologically equivalent species can coexist in drifting relative abundance for long periods of time with no refuge from intense diffuse competition. We can therefore no longer argue that species persistence is sufficient evidence by itself that a community is competitively stabilized by niche differentiation." (Hubbell and Foster 1986, 320) They use the community drift model to show that the competitive exclusion principle - if two species share the same niche for long, one will exclude the other - is false, for the community drift model shows that functionally identical, that is neutral, species can coexist for hundreds or thousands of generations. And without the competitive exclusion principle, which has variously been called an axiom and a law (Weber 1999), the opposing deterministic, equilibrium view loses a crucial piece of their argument. They go on to give two other arguments from the genetics of the forest and data about invasions. However, strong conclusions about the nonequilibrium status of the BCI forest community depend upon dynamic data since it concerns the species turnover and response to disturbance, data they did not have at that point.

I am telling this new narrative of the origins of the neutral theory both for its own, historical sake, and also to reform how we think of the epistemology of the theory. Towards the epistemological aim, I draw three important conclusions from rereading the 1986 paper that challenge the epistemic picture I opened chapter 1, section 1.2 with. First, the view of the neutral theory as a formal model of SADs used to infer neutral mechanisms is very inadequate to account for the use and the origin of the neutral theory. In this crucial early paper, the community drift model is not even fit to the data. Second, the dominant objection to all versions and stages of the neutral theory has been that the assumption of neutrality is clearly false. Yet here we find Hubbell giving evidence completely independent of model-fit for trees being approximately functionally equivalent. And third, this paper marks the beginning of the association between the neutral theory and the data from BCI. This connection must be made when telling the origins of the neutral theory.

2.4 The connection between the Barro Colorado Island Forest Dynamics Plot and the Origins of the Neutral Theory

Hubbell and Foster's 1986 paper is the first paper to use both Hubbell's work on the community drift model and data from Barro Colorado Island (BCI). Almost every later paper that Hubbell writes about the neutral theory will draw on the continually-growing BCI data. Therefore, it is worth probing into the connection between the neutral theory and the BCI.

Hubbell and Foster established the Forest Dynamics Plot (FDP) project on BCI in 1980. It is coincident, give or take a year, with Hubbell's 1979 paper and the first

appearance of the community drift model. BCI was the first field station of the Smithsonian Tropical Research Institute (STRI), established after the island formed from the construction of the Panama Canal. The core of the FDP project has been to completely census a 50 ha plot of forest every 5 years and collect large amounts of spatial and temporal data, larger than anything available for communities of any type before this. An enormous amount of data, research, and publications has been generated and continues to be generated by the FDP.

Two connections can now be made. First, the testing of equilibrium and nonequilibrium hypotheses was one of the stated purposes of the FDP. The BCI was not independent of his community drift model; it was explicitly one of the nonequilibrium hypotheses that would be tested using the new data. Hence they are linked from the beginning by the equilibrium problem. Second, the work on the scientists leading the BCI FDP take a very ‘close view’ of tropical forest ecology and this informs and supports the neutral theory of ecology, though in complicated ways. This is unexpected given Hubbell’s origin story and the narrative I presented in chapter 1 about formal models of SADs. This is what is glossed over, in Hubbell’s origin story when he says to “However, the germ of the ideas presented here can be traced by my thirty-year-old fascination with the origin and maintenance of high tree species diversity in tropical forests, an interest that launched several large-scale studies of tropical forest diversity and dynamics in the late 1970s.” (Hubbell 2001, xi)

In 2004, *Tropical Forest Diversity and Dynamism* collected the combined knowledge learned from the global Tree Plot Network, of which the BCI FDP is the

common ancestor and flagship (Losos and Leigh 2004). In the second chapter, Hubbell gives an overview of the context in which Foster and he established the plot, how it has changed, and what they have learned (Hubbell 2004). Hubbell writes that the BCI FDP was founded to test two groups of hypotheses of about the maintenance of diversity in tropical forests. The first is the Janzen-Connell hypothesis that host-specific predators cause the thin and even distributions of trees. The second is the regeneration/niche-gap partition hypothesis that falling trees created many varied, small environments into which different species specialized in growing in. He chronicles what they learned about these hypotheses over the course of the 20 years and what future work could bring.

It is striking that Hubbell barely even mentions that he is also the originator of the neutral theory of ecology, and that these two projects of his have been growing up together. His image would have been accurate if the BCI data was only used to test the neutral theory as a formal model once the neutral theory was well developed. But we have just seen that this is not true. Leaving out species differences is not merely a modeling heuristic, but has independent and direct evidence from a close view of tropical forests.

Hubbell and Foster completed their first census in 1982. They planned to complete the second in 1985 and then every 5 years after that. In the grant application to the NSF that they wrote in 1982 for the continued funding of the project, Hubbell and Foster lay out the long-range objectives of the plot: (1) to obtain large temporal and spatial statistics of the tree community, (2) to document the changes in forest composition directly and so test the various equilibrium and nonequilibrium hypotheses, and (3) to provide maps of tree distributions to other researchers (Hubbell and Foster 1982).

After discussing the various equilibrium hypotheses, they explain different nonequilibrium hypotheses and include, "... one of us (Hubbell) has proposed a model in which disturbance always reduces local diversity unless species are restored by immigration" and reference a paper in preparation by Hubbell called "Immigration, extinction, and the relative abundance of tree species in temperate and tropical forests".⁴⁵ Mention of Hubbell's model functions to transition to discussing the underappreciated role that chance and stochastic factors such as 'just being-there' play in a forest (Hubbell and Foster 1982). It goes on to argue that the important question (and this question is stable through all of Hubbell's papers on the neutral theory) is the relative significance of chance and determinism in structuring a tropical forest. The proposal explains that with better data provided by the BCI FDP, quantitative estimates of these relative strengths can be made. It is crystal clear from reading this and Hubbell's other papers that his ideas about tropical forest dynamics and composition and his ideas about the neutral theory are interrelated and mutually informing. Neither was the BCI FDP founded only to test Hubbell's hypotheses, nor was the neutral theory either an induction from new data or created to put new data to use. But all work on the neutral theory after 1982 will cite the BCI, and a significant amount of work of Hubbell's on the BCI will use the neutral theory in some capacity.

2.5 Conclusion

The connection to the BCI FDP further establishes that Hubbell's neutral theory is

⁴⁵ This paper was supposed to be presented at a symposium but Hubbell submitted a different paper. Hubbell says that it was never written. (Personal communication, 10/2012)

not a product wholly of taking a *distant* view to community ecology. David Lack looked *closely* at the island biogeography of Jamaican birds, and Hubbell looks *closely* at tropical forest ecology of the BCI and his theory reflects it. It does not *wholly* reflect it in that the community drift model was not an induction from BCI data. But they share a common origin in the equilibrium vs nonequilibrium problem. This connection counters the view the neutral theory of ecology as an abstract product of Hubbell's distant view of community ecology given in the narrative about formal model of SADs. It also supplements the way Hubbell's origin story glosses over the 'previous 30 years' and its association with the *distant* theory of island biogeography.

3. Why did Hubbell call his theory “neutral”? The neutral-null connection

I have shown how the history of the neutral theory of ecology can be placed in two distinct narratives within community ecology. The first narrative follows the problem of general mathematical modeling of the patterns and processes of biodiversity. Here the neutral model can fit patterns with the best of them and does so with an interpretation in terms of demographic and dispersal processes that ignore all interspecific competition. The second narrative follows the problem of determining whether communities are in taxonomic equilibrium or nonequilibrium. Here the neutral theory has been used as a nonequilibrium model of the state of tropical forests and coral reefs.

Hubbell told the origin story a bit differently in a 2008 paper than he did in the 2001 book.

“Nearly 30 years ago, I asked the question, what would the patterns of relative tree species abundance in closed-canopy forests be like if they were determined purely by demographic stochasticity in birth, death, and dispersal rates (Hubbell 1979)? This paper, my first foray into neutral theory, was a study of a tropical dry forest in Costa Rica. It unfortunately appeared during the height of the wars over null community assembly rules (Strong et al. 1984), ideas that were a logical outgrowth of the theory of island biogeography (MacArthur and Wilson 1967). The sacred cow in question was then, as now, the hypothesis that ecological communities are “niche-assembled,” that is, limited-membership, equilibrium assemblies of niche-differentiated species, each the best competitor in its own niche, coexisting with the other species in competitive equipoise. These wars were sufficiently off-putting that they delayed any further serious discussion of neutral theory in ecology for nearly 20 years. In my own case, I revisited neutral theory only in the mid-1990s, when a student in my Princeton biogeography class asked, why doesn’t the theory of island biogeography include a process of speciation, and what would happen if it did? I did not know, and I set about finding out.” (Hubbell 2008, 144)

This version of the origin story also mentions his prior work on tropical forests and the theory of island biogeography. But it foregrounds the connection between the community drift model and the ‘null model wars’, something he never mentioned in the book. The null model wars are given as the reason for the absence of work developing the community drift model further. The null model wars then set up the question, now asked by a student, of adding speciation to the theory of island biogeography.

The neutral theory is not described by Hubbell as “neutral” until the book is published in 2001. In this section I inquire into why Hubbell named his theory “neutral”. This inquiry shows how the origin of the neutral theory is related to the phenomenon of null and neutral modeling across biology.

I begin section 3.1 by showing the role that Graham Bell had on Hubbell's work. I then show how Bell introduced a distinction between the strong and weak versions or uses of the "neutral" theory of ecology. In section 3.2 I use this distinction to show how two other neutral theories of biology - Kimura's neutral theory of molecular evolution and the Woods Hole Group's MBL model of paleobiology - took differing stances on the strong vs weak distinction.

The rest of the section is devoted to building up to an understanding of how three distinctions fit together - strong vs weak, null modeling vs null hypothesis testing, and statistical null hypotheses vs dynamical null hypotheses. In section 3.3 I examine the beginning of the Null Model Wars in ecology between Daniel Simberloff and Jared Diamond, and use this to draw out the distinction between null modeling and null hypothesis testing. I then explain the other distinction that Bell introduced - statistical null hypotheses and dynamical null hypotheses. In section 3.4 I present my taxonomy of how these three distinctions are related. I explain how this taxonomy helps account for the failure to appreciate the difference between using the neutral theory of ecology, or any neutral theory, as a null model and as a biological model.

3.1 Graham Bell's call for neutrality

When did Hubbell first describe his theory as "neutral"? Hubbell published two new papers augmenting the community drift model with speciation: "Towards a Theory of Biodiversity and Biogeography on Continuous Landscapes" (Hubbell 1995) and then "A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs" (Hubbell 1997). Both were published very obscurely.

In the 1997 paper, he listed the title of his book in press as *A unified theory of biogeography and relative species abundance*. Then the book was published with the title *The Unified Neutral Theory of Biodiversity and Biogeography* (Hubbell 2001) in the prestigious Princeton Monographs in Population Biology series (of which *The Theory of Island Biogeography* was volume 1.)⁴⁶ He had never used the word “neutral” to refer to his own work before this. What led him to make this change? And is it connected to the null model wars?

Graham Bell published two papers just before Hubbell’s book came out: “The Distribution of Abundance in Neutral Communities” (Bell 2000) and “Neutral Macroecology” (Bell 2001). This is not a case of simultaneous creation, but of an interplay and even publicizing. The models Bell works with do not include speciation and are not much different from Hubbell’s community drift model. Bell clearly acknowledges Hubbell’s theoretical work, but does not speak as a popularizer. Bell argues that a neutral model is much-needed in ecology and shows that his concern is primarily methodological. Both papers write in terms of “null hypotheses”, but this term is ambiguous in ways that I show in this section.

Bell’s main purpose in the two papers is to raise a particular methodological question. He says, “... the success of the NCM [neutral community model] in predicting the major patterns of abundance and diversity has profound consequences for community ecology. These depend on whether the ‘weak’ or the ‘strong’ version of the neutral theory is adopted.” (Bell 2001, 2418) This comes in the conclusion, after Bell has devoted the

⁴⁶ Hubbell notes and discusses the disparity between where the papers and the book were published as a consequence of the controversial nature of the neutral theory in (Hubbell 2008).

paper to arguing for the adequacy of the neutral community model as a formal model of biodiversity and biogeography patterns. Bell's most basic NCM assumes that drift and immigration mechanisms are acting and that all individuals within the community compete with each other as equals for limited space.

The distinction between the strong and the weak versions of the NCM affects how the NCM can be used to make inferences.⁴⁷ He says,

“The weak version recognizes that the NCM is capable of generating patterns that resemble those arising from survey data, without acknowledging that it correctly identifies the underlying mechanism responsible for generating these patterns. The role of the NCM is then restricted to providing the appropriate null hypothesis when evaluating patterns of abundance and diversity.”

(Bell 2001, 2418)

The first way to interpret the ability of the NCM to fit observed patterns is to just acknowledge this ability without offering any explanation of it. In particular, no explanation in terms of the similarity between the assumptions of the model and the target system is offered. Its use here as “the appropriate null hypothesis” is as a check on alternative models which share its success but which are based on different ecological assumptions. For reasons I will explain more fully below, I prefer to call what a neutral model is being used as a “null model” instead of “null hypothesis”. The main difference is that null hypotheses are statistical hypotheses about the strength of a correlation between two variables, say smoking and cancer, while null models make substantial hypotheses about causes relevant to a correlation, say species differences and SADs. Also, I prefer to talk in terms of weak and strong ‘usage’ rather than ‘version’ because the difference is not in the content of the model (including both the mathematical model

⁴⁷ To my knowledge, Bell introduced this way of making the distinction.

and its ecological interpretation), but in the inferences made using the same model.

In virtue of being neutral, neutral models exclude species differences from having any effect on patterns. This means that they afford no causal responsibility to any processes such as interspecific competition or competition exclusion. Many competition-based models have been developed, and sometimes these are empirically adequate for describing and predicting some pattern. One explanation of this empirical adequacy is that, because the competition models work similarly to the communities being studied, the community is structured by interspecific competition. But when neutral models are used as null models with respect to competition models, then the empirical success (via model-fit) of a neutral model undermines the justification for the explanation of the pattern in terms of competition.

The *null modeling* reasoning strategy followed here is straightforward: The problem is deciding which theories or processes to accept and use to explain patterns of interest. If a neutral model is adequate to describe some patterns, then, because it is null with respect to species differences, species differences are not necessary to describe those patterns. Therefore, processes like interspecific competition which depend on species differences need not, and so further should not, be invoked. But this reasoning strategy cannot demonstrate the necessity of invoking dispersal processes or species similarities, because the neutral model is not itself tested. Furthermore, there is not a linear hierarchy of processes whose order must be followed; there is no reason why a neutral model must invoke speciation, or why a non-neutral model cannot invoke immigration. Therefore, it provides no confirmation of neutral models, and supports no causal explanation of patterns in terms of random drift, immigration, and speciation. The function of a weak

neutral model is to test particular competition hypotheses.

Contrast this weak usage with the strong:

“The strong version is that the NCM is so successful precisely because it has correctly identified the principal mechanism underlying patterns of abundance and diversity.” (Bell 2001, 2418)

Here the adequacy of the model-fits between the neutral model and the diversity patterns is evidence for the dominance of neutral or demographic/dispersal processes causing the patterns. Here similarity between model outputs and the observed distributions is taken as positive evidence for the similarity of the model and target system.⁴⁸ Bell goes on to explain that,

“This has much more revolutionary consequences, because it involves accepting that neutral theory will provide a new conceptual foundation for community ecology and therefore for its applied arm, conservation biology. We shall have, for the first time, a general explanation for community composition and dynamics, as well as a synthetic account of a range of seemingly disparate phenomena.” (Bell 2001, 2418)

Sometimes when someone says that a neutral theory provides the “null hypothesis”, they will say that it is because “the neutral theory supplies the baseline” or “is a baseline model” of the phenomenon. *Baseline modeling* is a reasoning strategy distinct from null modeling, but is superficially similar. Baseline modeling assumes that the neutral processes are going on in a community and that the strength of non-neutral processes like interspecific competition can be quantified through measuring the degree to which neutral models fail to fit observed patterns. Any deviation from the prediction of a baseline model are explained by additional causal factors not included in the baseline. And the lack of deviation from the predictions of the baseline model are explained using

⁴⁸ This is how (Giere 1998) would describe the basic model inference.

the neutral model. This is the crucial difference between null modeling, which *tests* hypotheses about causes, and baseline modeling, which *trusts* hypotheses about causes and measures their relative significance. Doing this relies upon accepting the strong version of the neutral theory. Hubbell does this both explicitly and implicitly, but not usually both together, and so it is important to keep in mind. The weak use excludes baseline modeling because the weak use of the neutral model does not warrant using the neutral model itself to give causal explanations.

If the NCM is treated as an explanatory model and used to infer the strength of neutral and dispersal causal factors, then it can serve as a conceptual foundation for community ecology and conservation biology in the sense of being a framework within which to work on conceptualizing, explicating, and measuring communities.

Nothing precludes the usage of the neutral model *alternatively* as an explanatory model and as a null model. But neither usage is implied by the other. The strong usage is stronger than the weak usage in the sense that the strong usage warrants using the neutral theory to give explanations. But the strong usage does not in itself justify the asymmetrical preference to neutral theory over competition theory, which is a respect in which the weak usage stronger.⁴⁹ This is subtle because the strong usage actually includes two parts: (A) that the neutral theory itself is treated as explanatory, and (B) that the neutral theory is the *correct* explanatory theory. Adopting part A of the strong usage alone only justifies the neutral theory as one possible general theory, on a par with all other theories, including competition theories, subject to empirical and theoretical

⁴⁹ The very fact that neutral hypothesis should be the ‘null’ hasn’t always been accepted. As I explain in section 3.2, selection and competition used to be regarded as the ‘null’.

support. Part B depends on empirical and theoretical support that is still open. Further, adopting both parts A and B do not require using the baseline modeling reasoning strategy identified above, which is itself controversial and a substantial methodological usage that requires independent justification. The mistaken belief that the strong usage includes the weak usage is partly responsible for why these two uses are sometimes run together and the neutral theory of ecology is used simultaneously as a null and explanatory model.⁵⁰

What is Bell's relationship to Hubbell and the neutral theory of the 2001 book? Bell begins his papers by noting the lack of any neutral or null models at use within ecology. He cites their presence within evolution with Kimura's neutral theory of molecular evolution and one dead-end attempt to bring this to ecology - Caswell's attempt in 1976. Hubbell cited Caswell as proposing the first *dynamic* model of diversity patterns in that the model is interpreted in terms of ecological mechanisms like birth and death and immigration in his first, 1979 paper. Hubbell's adaptation of Caswell's models however produced log-normal-like distributions which fit much better than Caswell's. Outstandingly, in the 1979 paper, Hubbell never even mentions the primary use which Caswell put the neutral models to - as a model which was explicitly null with respect to the causal factors assumed to be responsible by the leading account of the day - competitive exclusion.

It is interesting to track Hubbell's usage of "null" and "neutral" in his publications. Hubbell's first used "null" is in his paper with Foster in 1986 to describe the

⁵⁰ I argue for this directly in chapter 4.

community drift model,

“Although it is a “null” model, it should be noted that biotic interactions via diffuse competition (the sum of interspecific competition) are extremely intense in the model: No species can increase except at the expense of another species, since there is a finite limit on the total individuals of all species, since there is a finite amount in the world.” (Hubbell and Foster 1986, 320)

I interpret this to mean that the model is not a nothing-going-on null hypothesis. My evidence for this is that Hubbell gives a very similar statement in the book about the sense of “neutral” not meaning “nothing going on” but rather competition amongst equivalent individuals (Hubbell 2001, 6)

Neither (Hubbell 1995) nor (Hubbell 1997) ever uses “neutral” or “null hypothesis”, except to reference (Caswell 1976). Bell published his two papers on “neutral” ecology in 2000 and 2001, and then Hubbell renamed the theory the “neutral” theory. The first time Hubbell calls his theory the “neutral theory” is in the book. And here he defines neutrality just like Bell does, the identity of all individuals in the community in terms of ecological properties.

In the early paper, Hubbell and Foster never drew attention to the neutrality assumption, though he did state it:

“Indeed, if the probability of success is essentially the same on a per capita basis for all mature phase species, then the composition and relative abundance of tree species in the forest would tend to drift in a random walk.” (Hubbell and Foster 1986, 315)

Bell saw that, with Hubbell working on the community drift model again, a methodological discussion about the use of the neutral theory was called for. Bell saw that Hubbell’s work allowed the neutral theory to be used as both a null model and an explanatory model, and that these meant different paths for the neutral theory.

Hubbell first explicitly discussed the neutral-null connection in 2005. Here Hubbell showed how the neutral theory was related to work done by Stephen Jay Gould in paleontology. Hubbell identified the MBL model (on which more in section 3.2) as “neutral” for the same reason that the theory of island biogeography was neutral - because they both treat lineages or species equivalently. Though he uses the terms, Hubbell devotes almost no discussion to what it means to be “null”. He says,

“The principal use of the neutral theory is to evaluate when, and to what degree, asymmetries among species are required to explain the assembly of observed ecological communities.”

(Hubbell 2005, 123)

He notes that this is the same use to which the MBL model was put.

Hubbell first discusses the neutral theory a null model in detail in 2006, where he explicitly addresses Bell’s strong vs. weak question, and sides with the strong interpretation.⁵¹ Here he writes,

“The value of posing this question rigorously [what do you get from a neutral model?] is that we obtain a quantitative null hypothesis against which to test when, to what extent, and for which species demographic differences among species are necessary to explain observed community patterns.” (Hubbell 2006, 1387)

Hubbell will invoke the neutral theory as the supplier of null hypotheses/models in many later papers. I do not know what was responsible for this shift at this time. While Hubbell always argued for the causal responsibility of random dispersal, the use of the community drift model and then the neutral model changed. The community drift model was used to provide a how-possible explanation of SADs that undermined equilibrium explanations

⁵¹ Leigh made this observation (Leigh 2007).

based on model-fit, while the neutral model was used as positive evidence for random demographic and dispersal processes dominating based on model-fit.

3.2 The Strong vs. Weak usage in the History of Neutrality in Evolution and Paleobiology

Sketching the broader historical ties between the neutral theory of ecology and the history of null and neutral models in biology and the associated methodological wars is illuminating because it contextualizes the origins of the neutral theory, and because it shows both the denotations and connotations of “neutrality”.⁵²

Hubbell and other neutralists in ecology must announce whether they are adopting the strong vs weak position in order to decide what to infer from adequate and inadequate fits of the neutral model and observed data. Hubbell’s theory is useful as, and is presented primarily as such by Hubbell post-1997, an *explanatory theory* in the sense of being a biological model interpreted in terms of causal processes and used to offer an explanation and factors, as required by the baseline modeling strategy. It is also useful as a *null model* in sense of being a model explicitly constructed to exclude a set of causal process or factors and used test for the necessity of invoking processes explain a pattern.

Hubbell uses his theory as an explanatory theory and a null model. This can be seen

⁵² As far as I know, no attempt has been made by historians of science to chart out this phenomenon across biology. Huss’s dissertation is the exception (Huss 2004), which is focused on the MBL model. Huss discusses the MBL model in the context of null hypothesis testing and null and neutral models. Huss explicitly identifies the MBL model as a “neutral” model, and cites (Hubbell 2001) for his account of neutrality as identical probabilities for different units (Huss 2004, 58). He also explicitly distinguishes neutrality being null, as I have followed him in defining, and which is how Raup of the MBL group defined neutrality. Huss also identifies Hubbell’s neutral theory as an ancestor of the MBL model and a wider family of neutral and null models in biology (Huss 2004, 85, 138, 2009)

seen from his endorsement of the strong version in (Hubbell 2006). He consistently gives the neutral model a biological interpretation and never as a nothing-going-on, statistical interpretation. And yet he also praises its usefulness as providing a much needed null model for checking explanations of order in species diversity patterns in terms of selection and species differences. Null modeling and neutral theories have a history in biology, and Hubbell and other's use of the theory is conditioned by this history. Despite how chaotic the landscape around the neutral theory of ecology seems, it is a manicured park compared to the wild situation in the 1970s. There are now fairly delineated, if overgrown, accepted ways of thinking about the use of a neutral theory that emerged out of various episodes. To begin, I contrast the way in which the other two best known examples of neutral theories in biology were presented.

Motoo Kimura first developed formal models of drift in population genetics for their simplicity and mathematical tractability, not their accurate representation of real systems of alleles. Millstein, Skipper, and Dietrich have explained how James Crow and Kimura's 1970 *An Introduction to Population Genetics Theory* presents drift models as simplified first models to which complexity is added over the course of the work. These formal models of drift later were treated as accurate representation of the processes structuring real systems of alleles (Millstein, Skipper Jr, and Dietrich 2009).

Kimura first argued for the hypothesis that most variation in the frequencies of alleles in DNA are selectively neutral in (Kimura 1968). He did so with a comparative analysis of hemoglobin molecules in different animals to measure the rate of substitutions in chains of amino-acids. This is meant to give a good measure of the rate of evolution at

the genotypic level. Kimura shows that the rate of evolution is much larger than that expected when substitutions have high selective advantages and then argues that this is best reconciled by assuming that most substitutions are neutral or nearly neutral, meaning that most substitutions do not differ very much in fitness. He concludes,

“Finally, if my chief conclusion is correct, and if the neutral or nearly neutral mutation is being produced in each generation at a much higher rate than has been considered before, then we must recognize the great importance of random genetic drift due to finite population number in forming the genetic structure of biological populations.” (Kimura 1968, 626)

This argument begets the neutralist-selectionist controversy which still sees many publications to this day. Kimura and Ohta characterize the debate as “... whether the majority of mutant substitutions that occur in the course of evolution are due to random genetic drift or due to positive (adaptive) Darwinian selection.” (Kimura and Ohta 1977, 386).

What is important for our purposes here is that neither Kimura’s use of drift models as mathematical examples, nor Kimura’s rhetoric for the importance of genetic drift in real populations, invoke drift as the *null hypothesis* for testing selection hypotheses. Kimura never adopts the corresponding weak version of the neutral theory of molecular evolution or its predecessors.

Now it is common to refer to the neutral theory of molecular evolution as the “null hypothesis” or “null model” meant in the sense of a *baseline model* in which drift is assumed to be happening and deviations from pure drift are assigned to selection. For example, Michael Dietrich explains how Robert Selander expresses this in 1985,

““All our work begins with tests of the null hypothesis that variation in allele frequencies generated by random drift is the primary cause of molecular evolutionary change” (Selander,

1985, p.87). Selander notes that beginning with a neutral null hypothesis does not exclude selection as a possibility or predispose him toward neutrality. He starts with neutrality because he prefers “to begin with the simplest model” because it allows him to determine “a baseline for further analysis and interpretation”(Selander, 1985, p.88).” (Dietrich 2008, 164)

But this is the strong usage, not the weak usage.⁵³ The first reference I can find to the neutral theory of molecular evolution or any of its predecessors being a “null hypothesis” or “null model” is his 1983 book, “In the terminology of statistics, we treat the neutral theory as the null hypothesis and test it against observations.” (Kimura 1983) And there he goes on to say that such testing is too simplistic to settle the neutralist-selectionist controversy.

Kimura’s basic model is an (time-independent) equilibrium model where random drift and mutations balance out to produce an equilibrium for allele frequencies. This is analogous to how the neutral theory of ecology produces an equilibrium for SADs, on the basis of assuming the neutrality of alleles, that different alleles have the equal chances of being copied faithfully and of mutating in.

The rhetoric and methodologies such formally-similar theories of relative frequency distributions at different levels are very different. Kimura proposed his neutrality hypothesis at a time when the panadaptationist paradigm was at full strength - when adaptations were the most interesting patterns in biology, and were explained by invoking natural selection. Much of the controversy surrounding Kimura’s hypothesis and theory were misplaced because Kimura’s hypothesis was about the neutrality of the

⁵³ I make this point for Hubbell’s neutral theory in chapter 4. One way that baseline modeling is distinct from null modeling is that baseline modeling is a method for solving a relative significance problem (Beatty 1987), which assumes that theory choice is over. But null modeling is a method for theory choice.

genotypic level and left alone the accepted view that patterns in phenotypes were the result of natural selection. Yet it still served to challenge and circumscribe the dominance of natural selection. And the neutral theory of molecular evolution did so as an explanatory theory whose descriptive adequacy was (argued to be) positive evidence for a causal conclusion - the strong usage.

Compare and contrast this with use of the MBL model in paleobiology. The MBL model was the joint production of a group who met at the Woods Hole Marine Biological Laboratory (MBL) in 1972.⁵⁴ In their first paper “Stochastic Models of Phylogeny and the Evolution of Diversity” (Raup et al. 1973), David M. Raup, Stephen Jay Gould, Thomas J. M. Schopf, and Daniel S. Simberloff found a research program based around developing a stochastic model of the diversity patterns of taxonomic groups found in the fossil record that explicitly does not include natural selection. Legend has it that the model was designed to “leave Darwin out of evolution” after including Darwin was proving too complicated (Huss 2009).

The MBL model targets the patterns of diversity of taxonomic groups found in the fossil record. The Woods Hole Group was especially interested in how some taxonomic groups have more species within them than others, how some persist for longer than others, and how some groups diversify early and go extinct suddenly while others diversify and die out gradually. The MBL model is explicitly based on equilibrium theories like the theory of island biogeography. The group wanted to discover whether the balancing act between immigration and extinction could yield equilibrium diversities

⁵⁴ See (Sepkoski 2012) for the most authoritative history of this group.

in evolutionary time.

The MBL model works like this: begin with one lineage. Every turn, there are three possibilities: (1) the lineage can persist to the next turn, (2) the lineage can persist and branch into a new lineage, (3) the lineage can go extinct. These chances are determined by three independent parameters which treat all the lineages equally. Hence the MBL model is neutral in the same way the TIB is neutral - they each treat their basic units as having identical biological chances. Both the MBL model and TIB assume neutrality at the level of the taxon or species, while Kimura's and Hubbell's neutral theories assume it at the level of individual alleles or organisms. All lineages follow the same rules every turn. The MBL model actually depicts the approach to equilibrium and the equilibrium state. Before equilibrium diversity is reached, the branching rate exceeds the extinction rate and it set equal to it once equilibrium is reached. They summarize this as, "The maintenance of an equilibrium diversity in the present work implies that an adaptive zone or a geographic area becomes saturated with taxa and remains in a dynamic equilibrium determined by the opposing forces of branching (speciation) and extinction" (Raup et al. 1973). Further papers elaborated the research program and developed the models.

The group is very clear about their goals. The first paper outlines three goals: First, they will test the possibility that some aspects of the evolutionary record can be simulated well by random or stochastic variables. They are working with statistical randomness, the chance of a coin flipping, and showing that random processes need not entail random patterns or disorder.

Second, they will use their models to determine which patterns found in the fossil

record require interpretations in terms of non-random, specific and possibly unique causes. They argue that patterns which can be fit well by their stochastic model do not require explanation in terms of the natural selection, which is a directional, non-random cause or any other directional causes then accepted by paleontologists. Here we see the use of the MBL as a *null model* to check the selectionist explanations.

Third, they lay out their research programme - the model will be developed and altered in order to achieve ever-more adequate fits to the real world. They speak of modifying the way the branching mechanism in the model works or even in adding deterministic, directional processes. Under this process of model development, “The power of the model will then be its ability to specify the minimum departure from randomness necessary to produce a satisfactory replica of the real world situation.” This suggests the use of the model as a *baseline model* for measuring the strength of selection as the deviation from randomness.

The MBL model is initially proposed as a null model of diversity patterns to check the spread of unwarranted explanations in terms of natural selection. It would be interesting to track how the MBL moves around with the spectrum between explanatory model and null model. The number of people working on the project especially leads me to expect variation in the expressed positions on this. In the (half of the) group’s second paper in 1974, Raup and Gould wrote of the first paper,

“Throughout the work, we did not attempt to replicate nature but rather to identify those aspects of actual phylogenies which behave as stochastic or random variables. (We did this, in part, to establish “criteria of subtraction” for the identification of phenomena in actual phylogenies that cannot be simulated by random processes.)” (Raup and Gould 1974 #305)

This is exactly a statement of the agnostic position expressed by Bell as the weak version

of the neutral model of ecology. Even though the ‘criteria of subtraction’ is what is followed when baseline modeling, they never make the corresponding inference and apportion relative causal responsibility. I interpret this as due to their initial hesitancy to trust the MBL as a faithful baseline model. It is rather a how-possible explanation, while baseline modeling makes on how-actual explanations. We see then that the two early and significant neutral models were used in very different ways - Kimura’s first took the strong use as an explanatory theory and the MBL group first took the weak use as a null model.⁵⁵

Before leaving the point, the question of the strong vs weak uses of neutral theories has a history. Here are the two positions clearly expressed. In the preface of the 1987 *Neutral Models in Biology*, Matthew H. Nitecki and Antoni Hoffman summarize the debate like this,

“There is, however, more than just one theoretical model that can account for the empirical data and, therefore, model assumptions must also be tested. It is at this point that neutral models inescapably enter the argument. For if a model could be proposed which accounts for the pattern and yet involves no elaborate theory at all, and which follows from the principle of randomness under only a minimum number of assumptions that are unquestionable, why search for more sophisticated theories? The criterion of parsimony suggests that the neutral model should then be accepted as the explanation. But there is virtually no pattern that could not be obtained by some stochastic process, more or less complex, elegant or even plausible. What is, then, the appropriate methodology of testing neutral models? What is their use in the life sciences?” (Nitecki and

⁵⁵ The reasons for this are probably very varied and involve the personalities of the actors. But another important type of difference was probably the fact that molecular evolution had very good and extensive data while paleobiology did not. The data of the

Hoffman 1987, preface)

While presented in a dialectical manner, Nitecki and Hoffman clearly see “neutral modeling”, which is here as synonymous as two indefinite terms can be with “null modeling”, as a legitimate extension of classical statistical hypothesis testing. For them the crucial issue is the actual statistical testing. They are interested in how to best set the rates of type I and type II errors in a way that does not lead to neutral theories being accepted for everything in biology.

Contrast this with a short letter in the scientific correspondence in *Nature* in 1985 Leigh by van Valen,⁵⁶

“In the past decade or so a subtle misuse of null hypotheses has become almost standard in ecology, biogeography, functional morphology and theoretical paleontology. ... The difficulty here comes when a null hypothesis is placed in a privileged position, to be accepted at least provisionally until disproved.” (Van Valen 1985, 230)

The problem for van Valen is the privileging of a null hypothesis in the sense that it must be rejected before the alternative is entertained, *and* that if it is not rejected it is accepted. This clearly powerful position to be in combines both the strong and weak uses of a theory. I think van Valen is concerned with pushing back against the influence of hypothesis testing per se and Popperianism in biology.⁵⁷

With the MBL model and Woods Hole group we begin to see the larger methodological trends in biology influencing ecology. These trends include: the drive

neutral theory, especially after the BCI and associated forest dynamics plots puts ecology in between the two other sciences of relative frequencies.

⁵⁶ Thanks to Chris Haufe for showing the letter to me.

⁵⁷ I take this up in chapters 4 and 5.

towards becoming a science of laws or nomothetic science,⁵⁸ siding with equilibrium and general approaches over historical and particular approaches, an emphasis on *testing hypotheses* as marking scientific rigor, and a distinct Popperian view that current confirmatory science needs to be replaced with bold hypotheses and attempts to falsify them. Some of these trends follow Robert MacArthur. In particular, MacArthur *Geographical Ecology* reads as a manifesto for those biologists who want to move beyond natural history's marveling at the diversity of nature and want to get to the processes responsible for patterns. MacArthur's work also downplayed the role of history through (time-independent) equilibrium explanations and emphasized testing hypotheses against each other (Kingsland 1995). But there is a distinct lack of Popperianism in MacArthur, and MacArthur never challenged the strength of interspecific competition in shaping ecological patterns. For MacArthur, the neutrality of the theory of island biogeography was an outlier which he later tried to reconcile. Daniel Simberloff acted as the carrier between paleobiology and ecology on all these critical methodological positions.

In the next section I discuss the null model wars and their influence on the neutral theory. I use this to draw out the distinction between null hypothesis testing and null modeling as it emerged historically. I then show that the neutral theory is useful as a null model, not statistical null hypothesis.

3.3 The Null Model Wars

⁵⁸ The distinction between nomothetic and ideographic methods is not the same as Lack's close and distant views. The two distinctions are something like general vs particular

Daniel Simberloff was E.O. Wilson's (co-author of the theory of island biogeography) graduate student and was called on board the Woods Hole Group because of his work experimentally testing and evaluating the theory of island biogeography with real systems (Huss 2009). Simberloff had led the experiments to observe the recolonization rates of islands fumigated with methyl bromide. Simberloff is who most people think of when they think of null modeling in ecology, in part because of his outspoken nature.

The null model wars began in 1979 when Simberloff and Connor challenged Jared Diamond's view that the patterns of birds on particular islands were the result of interspecific competition. Since 1970, Simberloff had challenged ecologists on the grounds that the patterns they were claiming to be revealing, such as that "the mean number of species per genus (S/G) on an island is usually lower than S/G for its presumed source area" (Simberloff 1970), were not in fact significantly different from those expected by pure chance. Over the 1980s, Simberloff and colleagues at Florida State University such as Don Strong would become known as the Tallahassee Mafia and would continue to push a methodological change that centered around *testing hypotheses*, especially using the full power of classical statistics and computer simulations, supported explicitly by Popperian falsificationism. Diamond eventually would not attend events with Simberloff and many have said that the wars extended well beyond reasonable debate to personal attacks.

In 1975, Diamond published the massive "Assembly of Species Communities"

explanations of things, vs deductive vs inductive theorizing. Lack is then inductive but still general.

(Diamond 1975) wherein he presented and analyzed the co-occurrence patterns for bird species in New Guinea and surrounding islands.⁵⁹ Diamond compiled a large number of what he called “incidence functions” for different bird species. An incidence function for a species *A* graphed the proportion of islands occupied by species *A* (*J*), against the average number of species found on an island of a particular size-class (*S*). This is easiest to grasp visually.

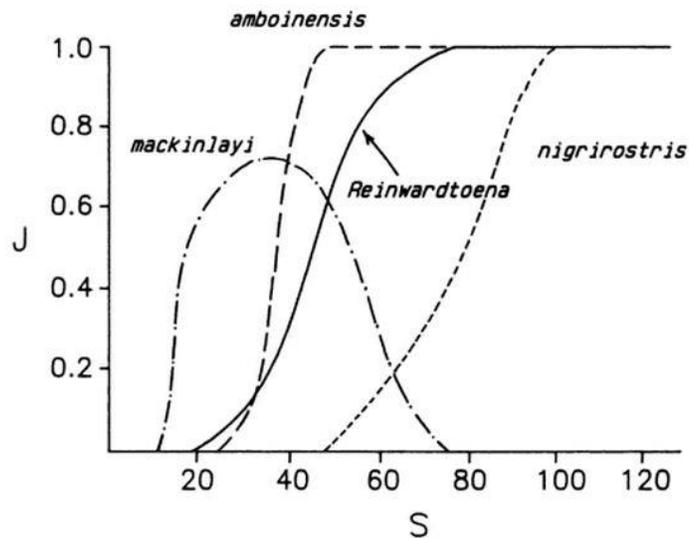


Fig. 4.2. Variation in the incidence functions of four species of cuckoo-doves on islands of the Bismarck archipelago. The functions plot the proportion of islands occupied by a species (*J*) against the mean number of species (*S*) characterizing a size-class of islands. The species *Macropygia mackinlayi* is a 'supertramp' confined to species-poor islands, whereas *M. nigrirostris* occurs only on species-rich islands. After Diamond (1975a).

Figure 2.1 – Incidence Functions of Cuckoo-doves on Bismarck Archipelago From (Wiens 1992, 80). Used with permission from Cambridge University Press.

Figure 2.1 compares the incidence functions of four related species on different islands. You read a particular incidence function, say for the *mackinlayi*, like this: one point is read as ‘the species *mackinlayi* inhabits *y*% of the islands with *x* many species on them’ and the whole line tells you about all the islands. Different species were said to fall into different categories: supertramps were found on only species-poor islands (low-*S*) and

⁵⁹ My discussion here is organized after reading (Wiens 1992), after reading the papers. 111

high-S species were found on all and only the species-rich islands.

Diamond used this organization of his data to examine the patterns of coincidence and exclusion between different sets of species. For example, for species within a guild of functionally similar species, some species, such as *mackinlayi* and *nigrirostris*, while they have overlapping incidence functions, they are never actually found on the same island together. Most dramatically, Diamond found that *mackinlayi* and *nigrirostris* never occurred on the same islands - this exclusive pattern was called a *checkerboard distribution*. In other checkerboard cases, combinations of species like AB, BC, and CD were found, but never ABC.

Diamond inducted from this data to seven empirical generalizations that he calls “assembly rules”.

1. If one considers all the combinations that can be formed from a group of related species, only certain ones of these combinations exist in nature.
2. Permissible combinations resist invaders that would transform them into forbidden combinations.
3. A combination that is stable on a large or species-rich island may be unstable on a small or species-poor island.
4. On a small or species-poor island, a combination may resist invaders that would be incorporated on a larger or more species-rich island.
5. Some pairs of species never coexist, either by themselves or as part of a larger combination.
6. Some pairs of species that form an unstable combination by themselves may form part of a stable larger combination.
7. Conversely, some combinations that are composed entirely of stable subcombinations are themselves unstable. (Diamond 1975, 423)

Diamond offered an explanation of these patterns. Diamond said that his

“working hypothesis is that, through diffuse competition [the sum of interspecific competition], the component species of a community are selected, and coadjusted in their niches and abundances, so as to fit with each other and to resist invaders.” (Diamond 1975, 343) To explain the assembly rules, Diamond assumed that the islands were already at a stable equilibrium and free from disturbances. For exclusion cases like *mackinlayi* and *nigrirostris*, he hypothesized that each species was better adapted to the particular conditions on the islands on which they are found than the other species. This led them to out compete and exclude the other species. Diamond acknowledged that chance sometimes moderated the determinism of competition through founder effects like changing which species got to an island first, but mostly chance played no significant role. For Diamond, interspecific competition was the most significant factor in determining species assembly on islands. This is in contrast to MacArthur and Wilson (MacArthur and Wilson 1967), who posited immigration and extinction rates as most important, and Lack (Lack 1976), who argued the *ecological poverty* of an island, its resources and areas, was the determining factor.

Connor and Simberloff challenged Diamond’s reasoning in 1979. Their abstract is worth quoting in full.

“We challenge Diamond’s (1975) idea that island species distributions are determined predominantly by competition as canonized by his “assembly rules.” We show that every assembly rule is either tautological, trivial, or a pattern expected were species distributed at random. In order to demonstrate that competition is responsible for the joint distributions of species, one would have to falsify a null hypothesis stating that the distributions are generated by the species randomly and individually colonizing an archipelago.” (Connor and Simberloff 1979, 1132)

They accuse Diamond of *assuming* the strength of interspecific competition in the communities and then explaining the assembly rules on that hypothesis. They charge Diamond with post facto rationalization and mere confirmation. To establish the hypothesis, Diamond needed to show that “a parsimonious null hypothesis” which hypothesized “no biological forces” were acting and which predicted biogeographical distributions failed to account for the patterns. They cite Popper’s *Conjectures and Refutations* (Popper 1963) to establish that science should proceed by “posing testable hypotheses and then attempting to falsify them”.

Our primary interest is in how Connor and Simberloff undermine the explanation of the assembly rules with competitive exclusion. They argue that the assembly rules would be expected if the species were distributed at random on the islands and that, since this hypothesis of randomness is the appropriate *null hypothesis* against which the competition hypothesis must be tested, the failure to reject the null hypothesis undermines the support for the alternative competition hypothesis.

Before I explain how Connor and Simberloff construct their null hypothesis, let’s walk through how the reasoning in standard classical statistical hypothesis testing works with the paradigm example.⁶⁰ Say a coin lands heads 30 times out of 100 and someone

⁶⁰ This is the way that Raup, fellow Woods Hole Group member with Simberloff, introduces the reasoning being followed when “neutral modeling” (Raup 1987, 121). Raup begins, “The neutral models used in paleobiology do not differ in any substantial way from the null hypotheses used standardly in applied statistics. Consider the classic case of coin flipping. ... The neutral models described in this chapter differ from the coin-flipping case only in being more precisely tailored to the biological problems at hand.”

I think he underestimates how different neutral or null models are from statistical null hypotheses. But, at a certain level of description, the reasoning pattern he identifies is shared across them all.

(Huss 2004, 117) discussed Connor and Simberloff’s reasoning with respect to Fisher’s example of the Lady Tasting Tea and null hypotheses.

argues that the best explanation for this was that the coin flipping was biased (for whatever reason, the coin itself, the flipping procedure, etc.). One way to challenge this conclusion is to assert that before any bias should be inferred, the null hypothesis that the coin flipping is actually fair needs to be rejected. One way to do this is to compare the likelihood of the observed frequency of heads to tails given the hypothesis of a fair flipping against the likelihood of the observed frequency given the hypothesis of a biased procedure. The null hypothesis of a fair flipping is that the probability of heads in a flip is 50%. The probability that a coin would come up heads 30 times out of 100 trials in a fair procedure is computed using the binomial distribution. Depending on how likely this is, how big the difference between the expectation for a fair coin and for the actually observed sequence, and the standards set before the test for how unlikely the null hypothesis has to be in order to reject it, a judgment is issued on whether to reject or accept the null hypothesis.

The reasoning goes, to whatever standards are accepted, if the null hypothesis of fairness cannot be rejected then the alternative hypothesis of bias cannot be inferred. In this way the null and alternative hypotheses are treated asymmetrically - the null must be rejected before the alternative is accepted, but not vice versa. The idea here is that, if you cannot reject the null hypothesis, then there is not evidence for any pattern in the data that requires explanation. Failure to reject the null hypothesis is a negative result. The fair coin hypothesis plays the role of a control group in an experiment. Failure to distinguish the experimental group from the control does not tell you that there was no difference, but that no difference was detectable with the data. No correlation between the shape of the coin and the outcome of the flip was established.

Now let us compare this reasoning to Connor and Simberloff's challenge to Diamond. Compare Diamond's competition hypothesis to the biased coin hypothesis and Connor and Simberloff's null hypothesis to the fair coin hypothesis. Connor and Simberloff claim that, if it can be shown that the patterns would not be expected on a random process of putting species on islands, then the competition hypothesis can be entertained. This is a substantial principle that they claim is justified by Popperian Falsificationism and statistical techniques for testing hypotheses. Accepting this principle, the difficult issue for the scientist is to construct the appropriate null hypothesis which excludes all and only the effect being tested - in this case, interspecific competition. For Connor and Simberloff, the problem is to determine the null expectation for birds randomly (with respect to species) distributed on islands. Much of the back and forth of the Null Model Wars, in addition to arguing about whether ecologists needed to be 'testing null hypotheses', was in developing the appropriate null expectation that did not secretly include competitive effects.

Connor and Simberloff's tested Diamond's rules 1, 5, and 7 with the following hypothesis. Pairs and trios of species should be randomly distributed on islands subject to three constraints: (1) for each island, the number of species on an island in the model is that which is observed, (2) for each species, species occur on as many islands in the model as they were observed on, and (3) species are only placed on islands which fall within the size-class of islands on which they are observed. They argued that condition 3 respected the use of incidence functions to establish ranges by Diamond. Then they constructed a computer simulation to determine the expected patterns on the null hypothesis. The details of the computer simulation are described at the end of the paper

and would be elaborated and critiqued in the subsequent papers. In short, they represented the co-occurrence of species on islands through a grid filled in with 1s and 0s and then ‘randomized’ them through sorting procedures and through drawing numbers of random sequences, subject to the three constraints. They argued that the patterns of birds on islands codified in the assembly rules was expected given their null hypothesis, and that therefore Diamond’s explanation of them in terms of competition was unjustified.

This exchange has different lessons for different groups and several for linking the neutral theory of ecology into this dialectic. Some of the disagreements between Diamond and Simberloff were due them talking past each other. For example, Diamond argued that the probability of particular birds being found in a checkerboard pattern was very unlikely. Simberloff responded by arguing that, in a random distribution it would be more unlikely if no two species were every found in a checkerboard pattern. This is like the counterintuitive result that, while it is very unlikely for you to win the lottery twice, it is virtually guaranteed that someone at some point will when the lottery twice. Hence, it is important to state and respect the particular problem posed.

The major criticism of Connor and Simberloff was to the rhetoric of statistical testing of null hypotheses. Gilpin and Diamond critiqued this directly,

“... the term “null hypothesis” is a misnomer that we place in quotes throughout this chapter, for these models are not at all null with respect to the posited factor but contain it, lack the logical primacy claimed for them by Simberloff and colleagues, and are merely one more alternative competing hypothesis.” (Gilpin and Diamond 1984, 298)

“The real world is not organized in the simple hierarchical manner that Simberloff and colleagues naively assume, where the effects of obvious factors can be stripped away first, leaving naked a separate for examination all the consequences (if any exist) of the debated factor. The “null

models” that these authors offer are misnamed: they are simply one of many unsuccessful competing alternative hypotheses, one that explicitly stresses some factors and implicitly contains others, including competition. We believe that it is probably inherently impossible to construct a useful everything-except-competition “null hypothesis”, and that further efforts in this direction will only sow more confusion.” (Gilpin and Diamond 1984, 315)

The use of the rhetoric of testing null hypotheses was one of the main fronts of the null model wars. In virtue of playing the role of the ‘null’ hypothesis, these hypotheses are given ‘logical primacy’ in the sense that the null hypothesis must be rejected before the alternative hypothesis can be sustained, but not vice versa. The hypotheses are treated asymmetrically. Gilpin and Diamond raise two important and lasting criticisms of this line of attack.

First, the hypotheses playing the role of null hypotheses in practice were not statistical hypotheses, but substantial biological hypotheses. The statistical hypothesis in the coin flipping case is just the Bernoulli distribution with $p=0.5$. But there is a difference between randomness in patterns and random processes. What Connor and Simberloff constructed and claimed were hypotheses of pattern randomness were in fact process hypotheses *random with respect to* certain causal factors being tested, which do not necessarily produce random patterns. Hence there is no general nothing-going-on statistical null hypothesis that can be used to test for whether there is a pattern or correlation that calls for explanation. Rather, what Connor and Simberloff are slipping into is giving an alternative explanation of the patterns of coexistence Diamond identified. Because they are actually proposing causal alternative hypotheses, Gilpin and Diamond argue that the ‘null’ hypotheses should not be treated asymmetrically and there should be so special burden of disproving the ‘null’, all hypotheses should be evaluated

on equal footing. Connor and Simberloff, to their credit, deny that they are doing this, and never take their analyses to support the hypothesis that the coexistence patterns are due to chance. But later, the neutral theory of evolution, the MBL model, and the neutral theory of ecology each eventually fall into the trap of using the rhetoric of testing null hypotheses to give explanations in terms of the hypothesis of chance and history.

What this shows is that there is another method being used here, which is falsely going by the name of ‘statistical null hypothesis testing’.⁶¹ I refer to this as *null modeling*. Null modeling accepts that a hypothesis which says that interspecific competition is responsible must show, not only that interspecific competition is a sufficient explanation for the data, but also that it is a necessary explanation. As van Valen pointed out, the main problem comes when null models are privileged as *accepted until rejected*. If failing to reject the null model is taken as a *negative* result, as it should be, some of their worry about the logical primacy of null models is assuaged.

Second, given that null modeling is using biological hypotheses to determine whether a cause, such as interspecific competition, should be invoked, it is extremely difficult to construct a null model that excludes all and only interspecific competition. The previous criticism was to all null modeling. This criticism is directed at the effectiveness in practice of null modeling. It is just not clear how, and may be impossible, to construct a model to test the role of competition on islands which removes all and only traces of interspecific competition and which would be rejected in the presence of all and only competition. For example, Diamond argues that the two constraints on Connor and

⁶¹ This point was argued at the time, very persuasively, by (Quinn and Dunham 1983, Sloep 1986). Nevertheless, as I argue throughout this dissertation, this knowledge was either not accepted or not retained in the case of the neutral theory of ecology.

Simberloff's null model which preserve the number of species on each island and number of islands inhabited by each species *include* the effects of competition. If the incidence functions result from competition, then a randomization procedure which respects these functions does not test for competition. This opens up the question of the constraints on a null model testing for some effect in a biological system. Do the number of islands need to be kept constant? In this case, yes. But clearly then different archipelagos will vary these. What about testing for the role of resource partitioning in determining the size ratios of different species? What about testing for the role of developmental constraints on the observed body plans? What do you randomize and what do you hold fixed? What about testing for the role of competition in determining the relative species abundances in tropical forests? This particular problem reinforces the general worry that these 'null hypotheses' for biological processes are not statistical hypotheses, but rather biology hypotheses, and should not be privileged as statistical hypotheses. This issue is acknowledged by null modelers as the prime difficulty invoked when null modeling by its practitioners to this day.⁶²

3.4 Organizing distinctions

We now have two distinctions to organize: the strong vs weak use of a theory, and statistical null hypothesis testing vs null modeling. When Bell raised these issues and applied them to the community drift model/ unified theory, he made one additional distinction that should be explained. Bell distinguished between *statistical null*

⁶² This difficulty is also shown in (Raup 1987) Gotelli and Graves's book is solely devoted to showing that this can sometimes be done and to getting better at null modeling (Gotelli and McGill 2006).

hypotheses and *dynamical null hypotheses*.⁶³ Following out this distinction will bring us full circle, back to the distinction between the weak and strong versions of the neutral theory qua null hypothesis.

Bell says, “Dynamic null hypotheses, such as the NCM [neutral community model], assert that the input to the system varies stochastically, whereas statistical null hypotheses assert that the output varies stochastically.” (Bell 2000, 614) Bell explains this distinction as the difference between a random pattern and a random process. But it is better illustrated as the difference between patterns which are produced by an independent process and by a dependent process. Take the flipping of a fair coin again. Whether the coin will come up heads on the 5th toss is independent of whether it will come up heads on the 3rd toss. The law of large numbers states that the frequency of heads of a fair coin will probably be around 50% for large numbers of tosses. But now consider that a body moves 1 step along a line in the positive direction when the coin lands heads and the body moves 1 step along the line in the negative direction when the coin lands tails. The outcomes of the coin tosses are still independent, but the motion of the body is not. Where the body will be found on the 5th turn is dependent upon where it is found on the 3rd toss. This is called a *random walk*. One property of a random walk is that there is no tendency for the walk to center around the origin; if a body moves away from the origin n steps in t turns, it is just as likely to be $2*n$ steps away in $2*t$ turns as it is 0 steps away. In this way, a random walk is dependent upon chance and history.

For ecological patterns of abundance and diversity, Bell says that the simplest statistical null hypothesis useful is one with no temporal or spatial dependence in a

⁶³ For the time being, treat these as proper names, not types of ‘null hypotheses’.

community. In the case of an island and mainland/metacommunity system, the island pattern would not be structured differently than the mainland pattern and the island pattern at a time was not related to the island pattern at an earlier time.

In a dynamical null hypothesis, the abundances of species in space and time are determined not only by chance but by history, and the patterns are shaped by dispersal processes of drift (birth and death), immigration, and speciation. The neutral model would be less structured if there were no immigration and speciation processes included and the community just drifted like alleles do in Hardy-Weinberg Equilibrium. Instead of the probability of a death being replaced by a given species being proportional to the current relative abundance of that species as it is in the neutral model, that probability could instead always be the same for all species or follow some other simple rule. This would erase the effects of history.

This partly explains why neutral models can produce structured patterns which cannot be distinguished from the structured patterns of models that incorporate species differences. The main work developing the neutral models that Bell and Hubbell do is in getting structured patterns out of a model based on random drift. The neutral model has been used to test Diamond's assembly rules and their explanation in terms of species differences and interspecific competition. Ulrich, using new techniques and software developed since the 80s, investigated what sort of expectations about the co-occurrence of species the neutral model produced (Ulrich 2004). And he has shown that the neutral model produces non-random patterns. It produces patterns less random than Simberloff's null hypotheses and more like Diamond's assembly rules. This is why Bell characterizes the distinction as between random patterns and random processes - because random

processes do not produce random patterns.

If being based on dispersal processes makes a neutral model a dynamical null hypothesis, what role does *neutrality* play? To answer this, we need to see that there are two ways to undermine an explanation of a pattern in terms of natural selection. One way is to show that there is no pattern in the data to be explained because the data are expected on null hypothesis. And here the null hypothesis could be statistical or dynamical, depending on the data. For example, further work by the Woods Hole Group, especially Raup, used random walk models to argue that evolutionary stasis was a pattern that required explanation. Rather than the statistical null hypothesis being no change in evolutionary rate over time, it became accepted that the appropriate null hypothesis was a random walk, or a dynamical null hypothesis. The neutral theory of ecology can serve this role qua *demographic* model based on drift.

The other way to undermine such an explanation is to argue that processes other than natural selection caused the pattern in the data. This cannot be achieved by a null hypothesis because the appropriate null hypothesis has already been rejected. Null models are used for this purpose. Null models are causal models null with respect to a set of causes used to produce expectations. Null modeling tests for whether a set of causes is necessary to produce a pattern. The neutral theory of ecology is used as a null model for interspecific competition in virtue of assuming *neutrality*. The neutral theory is null with respect to species differences and so can be used to test the necessity of invoking them as a causal explanation of patterns of abundance and diversity.

This brings us full circle back to Bell's strong vs weak dilemma. The weak usage

of the neutral theory serves only as a null to check the spread of rampant adaptationism and confirmatory science in community ecology. And the strong version of the neutral theory serves as a positive ecological theory. Table 2.1 below shows how these distinctions relate to each other. There are six possible ways for a theory to be used. Only two of them are satisfiable by the neutral theory of ecology; these are listed as “Open for NTE”, the others listed as “N/A for NTE”. There is some theory or hypothesis which is usable by each though. I also include whether drift is considered a causal process or a statistical idea, which I explain below.

Table 2.1 – Relationship between three sets of distinctions: Strong vs Weak, Null Hypothesis Testing vs Null modeling vs baseline Modeling, and Statistical Null Hypothesis vs Dynamical Null Hypothesis

Strong usage	Weak usage		
Baseline modeling⁶⁴	Null modeling	Null hypothesis testing	
N/A for NTE	N/A for NTE	N/A for NTE	Statistical null hypotheses
Open for NTE	Open for NTE	N/A for NTE ⁶⁵	Dynamical null hypotheses
<i>Causal drift</i>		<i>Statistical drift</i>	

This taxonomy shows that none of the distinctions exactly track each other. The four options given by null modeling vs null hypothesis testing and SNH vs DNH are all variations of the weak version of a theory. Being a dynamical null hypothesis does not

⁶⁴ Baseline modeling is not the only method that can be followed when adopting the strong usage, but it is the only one that I have considered in this chapter.

⁶⁵ The pure distribution produced by the neutral theory of ecology can be used as a statistical null hypothesis, but this is not the neutral theory.

require being used as a null model to test causal relevance. And being used as a null model does not require accepting the strong version of the model. The failure to reject either statistical or dynamical null hypothesis in either null hypothesis testing or null modeling never entails using a theory as a positive explanatory theory to give explanations and provide a framework. The failure to reject the null is always a negative result.

Nevertheless, being a dynamical null hypothesis opens the possibility of an interpretation in terms of biological processes. Drift has both a statistical and causal understanding.⁶⁶ I showed in chapter 1 that the same statistical distribution can, and often does, have a statistical and a biological interpretation. The shift from statistical understanding to biological understanding happens when moving from null hypothesis testing to null modeling and carries over into the strong version, which is explanatory. This explains, in part, why using the neutral theory as a null model and not a null hypothesis can be confused with using it as the strong version and not the weak version. The causal-statistical drift distinction does not track the strong-weak usage distinction. But using the neutral theory of ecology as a null model and failing to reject it as such in itself can only undermines support for interspecific competition.⁶⁷

⁶⁶ This distinction has been covered very well especially by Millstein for the neutral theory of molecular evolution and the MBL model (Millstein 2000, Millstein, Skipper Jr, and Dietrich 2009)

The primary difference between Millstein's work and mine is that Millstein is concerned to understand the nature of the chance in these theories, while I am concerned with how the concepts of chance are used in the reasoning and rhetoric. I completely agree with her that drift is a causal process, against the *statisticalists*, but I think that all three neutral theories are used in both statistical and causes senses and that this switching is interesting.

⁶⁷ I take up this difference in chapter 4 and argue analytically for why it is a mistake. Here I have shown why they should be seen as differently historically.

Hubbell is explicit that he takes the strong version and acts in accord with it when he makes causal conclusions from the fit between the neutral model and the BCI data such as “Species interactions, niche partitioning, or density-dependence, while they may be present, do not appear to enhance tree species richness at Barro Colorado.” (Condit, Chisholm, and Hubbell 2012b, 1), and “The neutral model predicts diversity and abundance at Barro Colorado because it properly describes what matters most – species input [neutral speciation] – while ignoring irrelevant details.” (Condit, Chisholm, and Hubbell 2012b, 5). In this way, Hubbell’s project is more similar to Kimura’s than the Woods Hole group’s. Hubbell drew on Caswell’s neutral models to move towards causal, dynamical, ecologically-interpretable models of abundance at the very beginning. But Caswell adopted the weak version of his neutral model despite drawing on Kimura’s strong version of the neutral theory of molecular evolution. Hubbell began by using the community drift model in the weak way but has used the neutral theory of ecology in the strong way.

4. Persuaded of these histories, what havoc must be make?

I have now presented three distinct narratives on the origins of the neutral theory of ecology. These narratives further interact in many ways. Together these give a different picture of the origins of the neutral theory of ecology than does Hubbell’s story. It is not surprising or interesting *that* Hubbell’s origin story is misleading. But a larger picture of the origins gives a different image of the science.

One reason that Hubbell's origin story gets told is because the theory of island biogeography is well respected in ecology and fecund across biology. The origin story serves to link the neutral theory to the theory of island biogeography in a way that diffuses some of the tension felt towards neutral theories. The theory of island biogeography was also neutral after all, and the neutral theory is just one, perhaps exceptional, of its many decedents. This shows *how* the misleading origin story is interesting.

But how do these narratives transform the philosophy of the neutral theory we began with? In the following five ways.

First, Neutrality is not the obviously false assumption it is made out to be. The BCI FDP shows that most individual trees are approximately neutral. There is evidence, completely independent from the neutral theory and model-fit, that shows that the BCI forest seems to be approximately neutral. The important question then becomes whether the BCI being approximately neutral justifies using the neutral theory as a *general* theory and applying it to other communities which are not approximately neutral.

Second, Neutral theory is not only the product of the abstract, distant view to ecology. It is connected to a particular close view, large-scale observational research program on the BCI. The connection between the neutral theory and the BCI FDP is most often seen as a weakness. Hubbell uses the BCI data in all the neutral theory papers, but even he does not emphasize the historical and ongoing connection between them. One explanation for this is that it is a minor instance of what we see in the historiography of the Copernican Revolution, where we traditionally celebrate Copernicus, Kepler, and Galileo for being theorists, and downplay the role of Tycho Brahe who finally did the

observations, without a telescope, which focused the new theorizing. And just as the dialectic next made Tycho the hero of the revolution and finally settled into giving each aspect its due, so to should we here recognize the importance and role of Hubbell and colleagues work to actually go and census a forest. I think that understanding the warrant for the neutral theory must take into account its connection to the BCI.

Third, characterizing the debate as competition vs neutral theory is very impoverished. The debate is also a relative significance controversy over whether competitive or neutral demographic and dispersal processes are more significant in particular communities and overall. But it not only a relative significance controversy either. The debate is neither only a competition vs neutral underdetermination problem nor only a competition and neutral relative significance problem. Both of these debates are going on together and they are interacting. This is the topic of chapter 4.

Fourth, the neutral theory is not limited to the problem of inferring processes from patterns such as SADs and SARs, but also pertains to whether communities are in equilibrium or nonequilibrium. Both the underdetermination problems and relative significance problems concern the equilibrium and nonequilibrium question too. It is not determinism *or* chance, but *how much* of each.

Finally, neutrality is not just *any* false assumption, but has a history and associated methodologies and inference patterns. Assuming neutrality is inextricably linked to claiming the theory as a null hypothesis, null model, and baseline model. These each come with their own research strategies, patterns of inference, and lines of argumentation. For example, claiming the neutral theory is a null model for testing competition theory with respect to SADs cannot license the conclusion that the neutral

theory accurately describes the SADs because random immigration is dominating there. This shifts from competition theory and neutral theory being exclusive alternatives to these theories being compatible explanations. I take this issue up in chapters 3 and 4.

Claiming the neutral theory is a baseline for measuring the strength of niche and neutral processes is interesting because of the way that the neutrality of the theory seems to be a reason for assigning deviations from the predictions of the neutral theory to niche theory. But what is the warrant for doing this? Moreover, if the neutral theory is assumed to be both an accurate description of processes going on in a community *and* often a poor description of these communities because other processes such as interspecific competition may be dominating them, why is the confirmation process for neutral theory the same as it is for statistical models like the lognormal distribution? I take these issues up in chapter 4 and 5.

Chapter 3.

Not Null Enough: Non-Statistical Null Hypotheses in Community Ecology and Comparative Psychology

Co-authored with Marta Halina

1. Introduction

A central goal in science is identifying and justifying the best explanation for a given phenomenon. There are many strategies for doing this, such as evaluating the empirical adequacy of the available hypotheses and assessing their relative epistemic virtues. Here, we examine a strategy for selecting between competing hypotheses used in the sciences. A central feature of this strategy is the use of what we call a “non-statistical null.” This is a hypothesis that practitioners in the field privilege as one would a statistical null hypothesis, but which lacks the features of a true statistical null. We present the ways in which non-statistical and statistical null hypotheses differ and argue that the privileging of non-statistical nulls lacks justification.

We begin in the second section by showing how the appeal to non-statistical nulls is used in community ecology and comparative psychology. In community ecology this strategy is used to defend the neutral theory of ecology, while in comparative psychology it is used to defend the behavior-reading account of social behavior in nonhuman animals. Both the proponents of neutral theory and the behavior-reading hypothesis characterize these hypotheses as “nulls” on the grounds that they are simpler than the available

alternatives. In the third section, we show that this form of argument is superficially similar to the statistical method of Neyman-Pearson testing, thus giving it rhetorical force, but go on to argue that it is disanalogous in precisely those respects required for justification. Whereas the null hypotheses used in Neyman-Pearson testing concern the effects of known, randomized extraneous variables within an experiment, the null hypotheses in our case studies are both general explanations for a phenomenon of interest requiring their own empirical support. We also extend our argument to counter the objection that model selection theory can be used to justify the non-statistical null strategy. We conclude in the fourth section that non-statistical nulls should be treated on a par with the available alternative hypotheses and the rhetoric of testing null hypotheses should be dropped. In the end, we hope our analysis will stimulate critical discussions on the role that nulls play more broadly in the sciences.

We are concerned with critiquing the non-statistical null hypothesis strategy, not the specific hypotheses advanced by this strategy. There may be better arguments or alternative lines of evidence favoring the neutral theory and behavior-reading hypothesis; we do not deny that this is the case. Our target is the strategy that depends on casting these hypotheses as nulls.

2. The Non-Statistical Null Hypothesis Strategy

The non-statistical null hypothesis strategy advances a particular explanatory hypothesis by appealing to it as a “null,” where its status as a null is justified on the grounds of simplicity. Rather than appealing to simplicity as an epistemic virtue, however, this approach grants the purported null the epistemic privileges of a statistical

null: it holds that (1) one must reject the null before one is warranted in accepting any alternative explanations and (2) if the null cannot be rejected, it should be accepted as the best explanation for the phenomenon under study. In what follows, we illustrate how this approach is used in community ecology and comparative psychology.

2.1. Explaining Relative Species Abundance Distributions in Community Ecology

A long-standing problem within ecology is to explain patterns of diversity and abundance of species within communities. A particular pattern of interest is the number of species found within a single trophic level and their population sizes—that is, distributions of relative species abundance. The traditional and still dominant approach to explaining this phenomenon holds that abundances result from competition between species for resources and trade-offs between how different species utilize those resources (Chase and Leibold 2003). For example, the well-known competitive exclusion principle holds that, if species coexist, then there are differences between how the species utilize resources. Mechanisms that depend upon species differences, including the mechanism of interspecific competition, are called selection mechanisms by analogy to natural selection.

Stephen Hubbell has challenged the selection hypothesis by advancing the neutral theory of ecology—a theory that is now both prominent and controversial in the discipline (Hubbell 2001, 2006). The neutral theory holds that communities are structured entirely by three mechanisms: ecological drift, random immigration, and random speciation. The theory is “neutral” because it assumes all individuals, regardless of species, have identical chances of giving birth, dying, immigrating, and being a member

of a new species. Because of neutrality, drift, immigration, and speciation are random with respect to species differences. The neutral theory predicts that a community's observed relative species abundance distributions will be a particular statistical distribution. The free parameters tuning the distribution are interpreted in terms of community size, metacommunity size, immigration rate, and speciation rate.

The neutral theory is explicitly constructed in order to exclude the influence of selection—for this reason it is said to be simpler than the selection hypothesis. Hubbell writes that the value of constructing the neutral theory of ecology is that, “we obtain a quantitative null hypothesis against which to test when, to what extent, and for which species demographic differences among species are necessary to explain observed community patterns” (Hubbell 2006, 1387). The neutral theory is not neutral in the sense of “nothing going on,” but rather in excluding species differences as causally responsible for the species abundance distributions observed in a particular community. This is the sense in which it is viewed as a null.

2.2. Explaining Chimpanzee Social Behavior in Comparative Psychology

Great apes such as chimpanzees engage in sophisticated social behavior. When stealing food from a competitor, for example, a chimpanzee will prefer to steal food that a competitor cannot see (Hare, Call, and Tomasello 2006, Melis, Call, and Tomasello 2006). The dominant explanation for such behavior is that chimpanzees track the mental states of others: they know (consciously or not) when other agents can or cannot see particular objects in their environment. Comparative psychologists currently hold that chimpanzees reason about a variety of mental states such as goals, intentions, and

perceptions (Call and Tomasello 2008). In accordance with human psychological research, this ability is referred to as “mindreading.”

Daniel Povinelli and colleagues have challenged the above consensus by arguing that there is a competing explanation for the apparent mindreading abilities of chimpanzees (Penn, Holyoak, and Povinelli 2008, Penn and Povinelli 2007, Povinelli and Vonk 2004, Penn and Povinelli 2009). Their “behavior-reading” hypothesis holds that chimpanzees employ a set of abstract behavioral rules that allow them to anticipate how other agents will behave on the basis of prior behavior and environmental circumstances. On this view, the fact that chimpanzees prefer to steal food that a competitor cannot see is best explained by the implementation of some rule, such as, ‘if there is an unobstructed line of gaze from an agent’s eye to a piece of food, then that agent will prevent the food from being stolen’.

Proponents of behavior reading refer to this hypothesis as a “null” on the grounds that it is simpler—in the sense of requiring fewer causal factors—than the mindreading hypothesis. On the mindreading account, a chimpanzee predicts the behavior of other agents on the basis of their mental states (e.g., they cannot see x), which is inferred from their observable behavior (such as the fact that their head is turned away from x). The behavior-reading hypothesis, in contrast, holds that chimpanzees do not perform the inferential step involving positing a mental state, but rather make their behavioral predictions on the basis of the observable behavior alone (Penn and Povinelli 2007, Lurz 2011).

2.3. The Asymmetrical Evaluation of Hypotheses

The neutral theory in community ecology and the behavior-reading hypothesis in comparative psychology are presented as null hypotheses by their proponents. The justification given for their status as nulls is that they are simpler than the competing alternative hypotheses because they appeal to fewer causal factors. Crucially, the treatment of these hypotheses as nulls drastically changes the way in which they are evaluated. Typically, two competing hypotheses are evaluated by comparing their empirical adequacy and theoretical virtues (such as internal consistency, coherence with other theories, generality, etc.). In contrast, the neutral theory and behavior-reading hypotheses enjoy a significant epistemic advantage. Proponents of these hypotheses hold that insofar as they are empirically adequate, they should be accepted as the best explanation of the phenomenon of interest. Unless the null can be shown to be empirically inadequate, all empirically adequate alternatives must be rejected regardless of their theoretical virtues. Anyone advocating for one of the competing alternative hypotheses must bear the burden of rejecting the null.

This asymmetrical treatment of hypotheses is exhibited in both of our case studies. In the case of community ecology, Hubbell maintains that, “obtaining acceptable fits from neutral models shifts the burden of proof to those who would assert that more complex theory is required to explain nature and with what level of detail and generality” (Hubbell 2006, 1387). Because the neutral theory is simpler than any theory that assumes species differences, the burden is on advocates of the more complex theory to disprove the neutral theory. If they are unable to do so, then one should conclude neutral mechanisms are dominant in producing the observed diversity patterns for that community.

Proponents of the behavior-reading hypothesis in comparative psychology also hold that it must be falsified before one is justified in accepting the claim that nonhuman animals mindread. As Penn and Povinelli write, “in order to produce experimental evidence for an f_{ToM} [theory of mind function], one must first falsify the null hypothesis that the agents in question are simply using their normal, first-person cognitive state variables.” (Penn and Povinelli 2007, 734) According to Povinelli and colleagues, until the behavior-reading hypothesis is rejected, we have “no evidence for theory of mind in animals” (Penn and Povinelli 2007, 732). Moreover, if the behavior-reading hypothesis cannot be rejected, we should accept it as the best explanation of nonhuman animal social behavior. Thus, Penn and Povinelli conclude, “the available evidence suggests that chimpanzees, corvids and all other non-human animals only form representations and reason about observable features, relations and states of affairs from their own cognitive perspective” (Penn and Povinelli 2007, 737).

3. Why Privileging Non-Statistical Nulls is Unjustified

The above asymmetrical treatment of hypotheses has important consequences on research. Rather than weighing the theoretical virtues of two empirically adequate hypotheses, the null hypothesis strategy advises one to accept the hypothesis taken to be the null while rejecting the alternative. In this section, we consider the justification for this move. We show that this approach mirrors Neyman-Pearson statistical hypothesis testing and consider whether it can be justified on the same grounds. We argue that it cannot be justified in this way because the purported nulls in our case studies lack the relevant features of statistical nulls.

3.1. Justification for the Asymmetrical Treatment of Hypotheses in Statistical Null Hypothesis Testing

Statistical hypothesis testing was developed in the 1920s and 30s most influentially by R. A. Fisher, Jerzy Neyman, and E. G. Pearson. Modern statistics has merged them and others into a toolbox of hybrid methods. Fisher and Neyman-Pearson testing differ with respect to both the number of hypotheses being considered in a test and the possible outcomes of that test. In Fisher testing, one hypothesis (the null) is tested and the possible outcomes of the test are either rejection of this null or failure of rejection. In Neyman-Pearson testing, two hypotheses (the null and the alternative) are tested and the possible outcomes are either acceptance of the alternative hypothesis (and rejection of the null) or retainment of the null hypothesis (and rejection of the alternative).⁶⁸

The strategy in our case studies mirrors Neyman-Pearson hypothesis testing in the number of hypotheses being tested, the possible outcomes of the test, and the asymmetrical treatment of hypotheses.⁶⁹ As Godfrey-Smith notes, the null hypothesis in Neyman-Pearson testing “typically gets the benefit of the doubt”(Godfrey-Smith 1994, 280). In this test, a type I error occurs when one rejects a true null hypothesis and accepts a false alternative hypothesis, while a type II error occurs when one rejects a true alternative and retains a false null. Typically, fewer type I errors are allowed than type II errors: whereas the accepted rate of type I errors is at or below 5%, the accepted rate of

⁶⁸ The language of “null hypothesis” comes from Fisher’s single hypothesis testing procedure, and Neyman and Pearson objected to its use, but it has long been a part of Neyman-Pearson testing. See (Gigerenzer 2004).

⁶⁹ Beatty (Beatty 1987) discusses the case of Kimura’s neutral theory of molecular evolution with respect to Fisher’s version of null hypothesis testing.

type II error is usually 20% and can be up to 30% or 40% (Sani and Todman 2008). In this way, the null hypothesis is privileged over the alternative. *Ceteris paribus*, it is better to accept the null over the alternative and better to reject the alternative over the null.

It is part of the standard practice of Neyman-Pearson testing to keep type I errors lower than type II errors. This raises the question of how one determines which hypothesis should be treated as the null and what justifies its privileged status. For Neyman and Pearson, the answer to this question was determined on pragmatic grounds. In discussing how one should go about making the decision of which error to avoid, they write:

We are reminded of the old problem considered by Laplace of the number of votes in a court of judges that should be needed to convict a prisoner. Is it more serious to convict an innocent man or to acquit a guilty? That will depend upon the consequences of the error; is the punishment death or fine; what is the danger to the community of released criminals; what are the current ethical views on punishment? From the point of view of mathematical theory all that we can do is to show how the risk of the errors may be controlled and minimised. (Neyman and Pearson 1933, 296)

Today, a more widely accepted justification for the asymmetrical treatment of error in Neyman-Pearson testing is that the null hypothesis is a hypothesis of “no effect” or “nothing going on.” Godfrey-Smith refers to this as the “semantic” justification because it concerns the content of the hypothesis, rather than the pragmatic implications of accepting it. Although Godfrey-Smith does not go into detail about what he takes the semantic justification to be, he holds it is grounded in the value of simplicity. He writes,

If hypotheses of “no effect” are nulls, then the asymmetry between alpha [the rate of type I errors] and beta [the rate of type II errors] operates as the wielder of Occam’s razor. The more serious error is multiplying effects beyond necessity, rather than not recognising enough effects. The asymmetry establishes a bias in favor of the simpler hypothesis. (Godfrey-Smith 1994, 282)

If one is uncertain about whether there is a relationship between two variables, it is best to err on the side of holding that there is no relationship. Godfrey-Smith is right here to point out that the null hypothesis is favored in statistical hypothesis testing because it is taken to be a hypothesis of no effect, where “no effect” means that the independent variable being tested did not affect the dependent variable being measured. However, as we argue below, it is a mistake to understand this favoring as simply the wielding of Occam’s razor. Understanding the asymmetrical acceptance of error in the context of statistical hypothesis testing requires situating it within the practices of experimental design and statistical inference.⁷⁰

The role of the null hypothesis in statistical hypothesis testing must be understood within the context of a well-designed experiment. The purpose of an experiment is generally to test whether one variable (the independent variable) has some effect on another (the dependent variable). Researchers do this by testing whether there is a difference in the dependent variable between conditions in which the independent variable is present (the experimental condition) and absent (the control condition). A well-designed experiment requires that one control all extraneous variables—i.e., those variables (other than the independent variable) that could or are known to have an effect on the dependent variable. Researchers control these variables by either eliminating them or distributing them randomly across conditions. For example, the age and education level of a participant may affect how he or she performs on a memory task. Knowing this, researchers randomly assign participants to the control and experimental conditions.

⁷⁰ For general introductions to statistical hypothesis testing see (Sani and Todman 2008, Dienes 2008).

Statistical inference is then used in order to determine whether these variables could have led to the observed experimental results.

Researchers use inferential statistics in order to determine the likelihood that randomized extraneous variables are responsible for the results of the experiment. A statistical null hypothesis is employed for this purpose. The *statistical null hypothesis* states that there is no significant difference in the mean outcome of the control and experimental conditions, where a significant difference is one that can reasonably be attributed to the independent variable. If researchers could eliminate all extraneous variables, and the independent variable failed to affect the dependent variable, then the results of the control and experimental conditions would be identical. In such a case, if the independent variable were not to produce an effect, there would be no other variable that could. A world without extraneous variables would be one in which factors like age and education have no effect on participants' performance on a memory task. In the real world, extraneous variables are always a source of variability. The question then is whether this variability is sufficient to account for the differences observed across conditions. If it can, then one accepts the null hypothesis that the observed difference is not significant—that is, not large enough to be attributed to the effects of the independent variable.

How does one determine whether an observed difference can be attributed to the independent variable? Roughly, one calculates a statistic that reflects the variability of the data within conditions, where we know the independent variable is not having an effect. For example, researchers can estimate the effects variables like age and education have on participants' performance on a memory task by looking at how their performance

varies within the control condition and within the experimental condition. The variability within these conditions should be attributable to the randomized variables of age and education level since all other variables within these conditions have been held constant. This variability or standard deviation is then used to calculate the probability that the difference observed between conditions is due to extraneous variables alone. Again, if this probability is small (traditionally, less than 1 in 20), then one concludes that the observed difference is significant or the independent variable likely had an effect.⁷¹ If this probability is large, on the other hand, then one concludes that the observed results were likely the product of the randomized variables. In the latter case, the experiment has not produced evidence for a relationship between the independent and dependent variables.

Now we can see more clearly why researchers err on the side of accepting the null hypothesis in statistical hypothesis testing. It is not because it is simpler to posit no relationship between the independent and dependent variables. Instead, the null hypothesis is favored because the values of the extraneous variables are known and have been shown not to correlate with the observed variation in the dependent variable. The effects of the extraneous variables are also of little interest: the purpose of the experiment is not to determine their effects. To retain the null hypothesis is to recognize that the observed results were likely caused by those variables the researchers did their best to eliminate. When there is a high probability that these extraneous variables (or “nuisance variables” as they are also called) are responsible for the results, one concludes that the

⁷¹ Neither accepting the experimental hypothesis or retaining the null is a conclusion of certainty. One decides to accept the experimental hypothesis because it is unlikely that the observed effect was caused by extraneous variables alone. Likewise, one decides to retain the null hypothesis not because we are certain the independent variable had no effect, but because the chances of extraneous variables producing the observed result are high.

independent variable likely did not have an effect or the present experiment at least was unable to detect it. In contrast, the effects of the independent variable are not known. Uncovering these effects is the purpose for which the experiment was designed. There may be other evidence suggesting that the independent variable should have an effect in this situation, but the aim of the study is to provide experimental evidence that this is the case.

The difference between retaining the null hypothesis versus the experimental hypothesis within the context of statistical hypothesis testing becomes apparent when one considers the fact that retaining the null is a negative result. Mayo and Spanos refer to the contrary view, that the failure to reject the null hypothesis is positive evidence of no effect, as the “fallacy of acceptance” (Mayo and Spanos 2006, 338). When researchers have failed to reject the null, they do not hold that they have produced evidence in its favor. Instead, they conclude that the experiment failed to provide evidence for the experimental hypothesis—either the experimental hypothesis is wrong or the experiment needs to be better designed (extraneous variables need to be better controlled, sample sizes need to be larger, etc.). Contrast this with Godfrey-Smith’s semantic justification of the null. Here the claim is that both the null and experimental hypotheses are positive, empirical hypotheses tested by the experiment. The main difference between them is that the null is simpler, positing that the independent variable will not have an effect. But, as we have seen, statistical-null hypotheses are not positive, empirical hypotheses. They represent experimental noise that is known and unwanted, but that researchers must take into account in order to test the experimental hypothesis of interest.

The non-statistical null hypothesis strategy is analogous to Neyman-Pearson

hypothesis testing in many respects. It relies on two hypotheses, a null and an alternative; it couples rejection of one hypothesis with retainment of the other; and it privileges the null hypothesis over the alternative. Given these similarities, one might think that this strategy is justified as an extension of statistical null hypothesis testing. In the next section, we argue that this is not the case. The non-statistical null hypothesis strategy differs from Neyman-Pearson testing in precisely those respects necessary for its justification.

3.2. Why Non-Statistical Nulls are Not True Nulls

The role statistical nulls play in hypothesis testing depends heavily on the context of experimental design and statistical inference in which they are used. The fact that researchers require that one reject the null before accepting the experimental hypothesis makes sense within this context. Retaining the null means the noise created by randomized extraneous variables was too large for the potential effects of the independent variable to be detected. It does not mean that one has tested the effects of one or more of these extraneous variables and found that they produced the experimental result. Retaining the null is a negative finding.

Proponents of neutral theory and the behavior-reading hypothesis take their hypotheses to be epistemically privileged in the same manner as a statistical null. However, these “nulls” lack the relevant features of a statistical null. First, they are not used within the context of a well-designed experiment to refer to extraneous variables that have been intentionally randomized by experimenters. Second, there is no evidence that the causes posited by these nulls are operating in the form of within-condition

variance. Instead, the behavior-reading and neutrality hypotheses refer to mechanisms or causal factors in the world that proponents are positing may be responsible for the phenomenon of interest, but which there is not yet evidence that this is the case. They are identical to the alternative hypotheses (selection theory and mindreading) in these respects; they are *alternative* alternative hypotheses, not nulls.

In the case of neutral theory, Hubbell defines what he intends by “neutral”, implicitly contrasting it with “null”. He writes:

Before proceeding, I need to be more precise about the meaning of *neutrality* as used in this book. Despite its moniker, the concept of neutrality actually has many meanings in the literature. To most people, the word *neutral* congers [*sic*] up the qualitative notion of “nothing going on.” But exactly what people mean by this phrase often turns out to differ from one person to the next. I use *neutral* to describe the assumption of per capita ecological equivalence of all individuals in a trophically defined community. This is a very unrestrictive and permissive definition of *neutrality* because it does not preclude interesting biology from happening or complex ecological interactions from taking place among individuals. (Hubbell 2001, 6-7) Emphasis in original.

The behavior-reading hypothesis is also a claim about a mechanism operating in the world. It holds that chimpanzees possess a set of behavioral rules or heuristics, which allow them to make predictions on the basis of observable situations. According to proponents of the behavior-reading hypothesis, the mindreading hypothesis makes a very different causal claim. Penn and Povinelli characterize the difference between nonhuman behavior-reading and human mindreading as follows:

we believe that human and nonhuman animals possess a variety of mechanisms for recognizing those relations that are causally relevant to predicting the goal-directed behavior of other intentional agents. These heuristics enable both human and nonhuman animals to pick out the causally relevant relations between ‘what’ an agent is ‘looking’ at and how that agent is likely to

behave in the near future... However, only humans cognize the higher-order analogical similarities between perceptually disparate behaviors and thus only humans possess the ability to reinterpret other agent's goal-directed relations in terms of abstract mental state relations disembodied from any particular task context. (Penn and Povinelli 2013, 20)

Penn and Povinelli take the mechanisms involved in behavior-reading to be causally distinct from those employed in mindreading. The hypothesis that nonhuman animals behavior-read is not merely the claim that they lack the ability to represent the mental states of others, but also the claim that they possess a set of heuristics, which allows them to make the correct behavioral predictions without these abilities. If the human ability to reinterpret behavior in terms of mental states had no causal effects, then this ability would be superfluous and impossible to test empirically. Povinelli and colleagues, however, do not think this is the case (Penn and Povinelli 2007).

For both the neutrality and behavior-reading hypotheses, the mechanisms posited are claims about some unknown aspect of the world, rather than the extraneous variables of a particular experiment that researchers have intentionally randomized. Moreover, proponents of these non-statistical nulls take the failure to reject these posited mechanisms as evidence that they are in fact operating in the world, but this commits the fallacy of acceptance (Mayo and Spanos 2006, 338). Retaining a statistical null, in contrast, amounts to no more than concluding that an experiment has failed to detect a relationship between an independent and dependent variable. The features justifying the privileged status of a statistical null are precisely those that are missing in the case of non-statistical nulls.

3.3. What about Model Selection Theory?

We have argued that non-statistical null hypotheses like the neutral theory and behavioral-reading hypothesis are relevantly disanalogous from statistical null hypotheses. Given this, they should not receive the epistemic privileges that they currently enjoy. One might object to this by arguing that this epistemic privilege comes not from an analogy to statistical null hypothesis testing, but from model selection theory.⁷² In this section, we argue that this is not the case.

The Akaike model selection criterion, as expounded by Forster and Sober (1994), gives a uniquely best way of balancing the respective virtues of the fit between a model and data set and the number of free parameters in that model. A model's number of free parameters is considered the measure of its simplicity in the Akaike framework. Akaike proved a mathematical theorem that provides a unique answer to whether the additional cost of added complexity in terms of the number of free parameters is worth the benefit of increased fit. Given a set of statistical models and a data set, the Akaike criterion selects the model that is likely to be most predictively accurate with respect to future data drawn from that data set. The Akaike criterion rewards models for their goodness-of-fit to the given data set and penalizes them for their number of free parameters. With too much goodness-of-fit or too few free parameters, more data will probably make the fit worse. Akaike's theorem gives a unique way of weighing a model's goodness-of-fit against its number of free parameters. The model with the smallest value is selected. In this way models are privileged for their simplicity.

Simplicity is traditionally considered to be a theoretical virtue, distinct from empirical virtues concerned with evidence. But Sober argues that the choice of a best

⁷² We thank Elliott Sober for raising this point.

model using the Akaike criterion is a purely empirical decision in the same way a test using only goodness-of-fit is empirical. He writes that, “although it is clear that simplicity is a separate consideration in model selection from fit to data, the justification provided by Akaike’s theorem for using simplicity depends on empirical assumptions. Simplicity is therefore an empirical consideration” (Sober 2002, S117). Given this, one might argue that the non-statistical nulls in our case studies should be privileged because their simplicity makes them more empirically successful than the alternatives. Although it may seem that the two hypotheses fit the data equally well, the non-statistical nulls are superior because they have fewer free parameters. Simplicity trumps the theoretical virtues held by the alternative hypothesis, such as coherence with other theories or generality, in the same way goodness-of-fit trumps these theoretical virtues.

There are two main objections to this strategy. First, as presented above, the non-statistical nulls in our case studies are viewed as simpler because they exclude a cause included in the alternative hypotheses. Neutral theory is simpler than selection theory because it does not posit selection acting in a community; behavior reading is simpler than mindreading because it does not posit that individuals represent mental states. However, as Sober argues, simplicity qua free parameters is not the same as simplicity qua excluding a cause (Sober 2009, 247). Free parameters are a feature of mathematical models and causes are features of theories, hypotheses, and interpretations of mathematical models. Two models might be interpreted in terms of the same causes and differ in their number of adjustable parameters—treating two causes as interactive rather than additive, for example, adds a free parameter. One model might also be interpreted as including fewer causes than another, while having more free parameters. For example,

the broken-stick model is a selection model with two free parameters: number of species and number of individuals (MacArthur 1957). However, one of the formalization of neutral theory includes three free parameters: local community size, immigration rate, and a function of metacommunity size and speciation rate. The latter is simpler than the former in the sense of excluding selection as a cause of species abundance distributions, but not in the sense required by the Akaike framework.⁷³ The non-statistical nulls in our case studies cannot be defended using model selection theory without first showing that they have fewer adjustable parameters than their competitors.

Current descriptive inadequacy does not necessarily disqualify the Akaike framework as a potential justification for non-statistical nulls. Perhaps these hypotheses can be revised in ways that will make them amenable to this method of analysis. Unfortunately, this route does not look promising for proponents of the neutral theory and behaviour-reading hypothesis. The Akaike criterion is a tool for comparing mathematical models on a very limited number of dimensions: fit with a data set and number of free parameters. When it comes to identifying and evaluating the causes responsible for a given distribution of data, the Akaike criterion is silent. The mismatch between the Akaike criterion and the criteria needed to evaluate causes becomes clear in Purves and Pacala's critique of the neutral theory. They constructed a selection model that was mathematically equivalent to a formalization of the neutral theory, and argued on these

⁷³ Stephen Hubbell will speak of the small number of free parameters in his model as a virtue. The earliest he does this is (Hubbell 2008). However, this is not the only or primary way in which he and others appeal to the simplicity of the neutral theory.

Einstein said the epistemology of a working scientist appears as "opportunistic" to the systematic epistemologist. We see Hubbell as an epistemic opportunist because he is pursuing multiple argumentation strategies simultaneously, using various tools to suit these ends.

grounds, that the two models were equally well supported (Purves and Pacala 2005). Hubbell rejects this conclusion, however, maintaining, as quoted above, that “obtaining acceptable fits from neutral models shifts the burden of proof to those who would assert that more complex theory is required to explain nature” (Hubbell 2006, 1387). Hubbell is not referring to the number of free parameters in the models here because the two models are equivalent. He means that the neutral theory should be favored on the grounds of its relative simplicity qua excluding selection. If proponents of non-statistical nulls move in the direction required to apply model selection theory, they may very well lose their privileged status as nulls. This is fine, but if they take this route, they cannot then appeal to criteria outside of this framework when things do not work out in their favor—at least not without justification.

Objections aside, applying the Akaike framework as a method for choosing among alternative hypotheses embodies scientific values that may not be shared by scientists, so such a move requires explicit discussion and justification. The Akaike criterion considers only statistical goodness-of-fit and number of free parameters, yet there are many other scientific virtues which should be considered when evaluating alternative hypotheses. In an earlier episode in biology, Roughgarden (Roughgarden 1983a) and Beatty both criticized the rhetoric of “testing hypotheses” for being used to support a particular form of statistical testing. They argued that, “the logic of hypothesis testing in question... [should not be] misinterpreted as providing simple answers to difficult problems about the priority of scientific pursuits” (Beatty 1987, 60). We agree. Evaluating scientific hypotheses is a difficult problem that cannot be decided by a mathematical theorem. One must show that this theorem is the right tool for the job,

given a community's research goals.

4. Parity Please

We have argued that the intended justification for privileging non-statistical nulls, as an instance or extension of statistical null hypothesis testing, is unjustified. Like selection theory and mindreading, neutral theory and the behavior-reading hypothesis are not statistical hypotheses. The appeal to these latter hypotheses as “nulls” is a rhetorical strategy that lacks justification. It presumes that the virtue of simplicity trumps all others without providing good reasons for it doing so. The use of the term “null hypothesis” serves to gloss over this in the name of good scientific, statistical methodology.

Richard Levins made the tradeoffs between scientific virtues among different models and modeling practices well known (Levins 1966). He argued that simplicity is one virtue among many that all exist in complex trade-off relationships. Which virtues should be valued over others depends on the problems, goals, and circumstances of the particular situation. We find such tradeoffs in our case studies. For example, species similarities and differences are used to explain a wide variety of empirical results, such as why a particular type of tree is found in the same temperature belt across continents. This result is not in the neutral theory's explanatory scope. Appealing to the neutral theory on the grounds of relative simplicity in order to explain some patterns of biodiversity and biogeography means researchers are choosing to value simplicity over generality. Instead of relative species abundance distributions and continental species patterns being explained in the same way, they would be explained with different theories. This may be a tradeoff researchers endorse, but it is a choice. Conceiving neutral theory as a non-

statistical null takes this choice away from researchers. It makes the favoring of simplicity uncontroversial and inevitable. Tradeoffs between virtues require choices, which in turn require reasons and justification that can be discussed, critiqued, and evaluated.⁷⁴

In addition to being weighed against other virtues, the value of simplicity must be assessed against the particular context in which it is employed. Sober (Sober 2002) and Norton (Norton 2003) have argued that simplicity is not a general virtue, but is, whenever justified, shorthand for some set of facts which make the simpler hypothesis more likely. As Sober writes,

When a scientist uses the idea [of parsimony], it has meaning only because it is embedded in a very specific context of inquiry. Only because of a set of background assumptions does parsimony connect with plausibility in a particular research problem. What makes parsimony reasonable in one context therefore may have nothing in common with why it matters in another. (Sober 1994, 140)

If Sober is right, then an appeal to simplicity must be evaluated on a case-by-case basis.⁷⁵

It may be that simplicity should trump other virtues, in the way that empirical adequacy and goodness-of-fit sometimes does. However, until an argument is provided for this, the value of simplicity should be treated on a par with other virtues. The non-statistical null hypothesis strategy is an attempt to move hypotheses away from parity by shifting the burden of disproof to the alternative hypotheses on the authority of statistics. As we have argued, there is no clear justification for this strategy, however, so the playing field should be leveled.

⁷⁴ Such a discussion between simplicity and other scientific virtues may be particular, as was (Levins 1966), or general, as for example (Douglas 2009).

5. Conclusion

The non-statistical null hypothesis strategy is employed on behalf of the neutral theory of ecology and behavior-reduction hypothesis in comparative psychology. Because the strategy presents a rhetorically powerful solution to a common problem, we expect these are not isolated cases. However, we have argued that this strategy is not justified. Hypotheses should not be privileged for being “nulls” when they are not statistical nulls unless a valid justification has been given for such privileging.

⁷⁵ (Longino 2008) also argues against simplicity having value in general.

Chapter 4.

Modeling: Neutral, Null, and Baseline

1. Introduction

Community ecologists want to understand and causally explain basic features of communities such as why the tropics have the greatest diversity, and why it is that, no matter the community, most species there will be rare and few will be common. The standard approach to these kinds of questions begins with an appreciation of the diverse ways that different species utilize resources and effect the environment. The hope is that studies of specific species will fit together like puzzle pieces and our understanding of ecological community will follow from our understanding of the competitive interactions between species. Call this approach *competition theory*. An alternative approach to these questions begins instead by disregarding all ecological differences between species in a community. Instead of competitive interactions, community patterns are the result of history and chance. Call this approach *neutral theory*.

In this chapter I show that a mistake is being made in the use of the neutral theory of ecology that undermines the causal conclusions being drawn. Ecologist Hubbell and others argue that because (1), the neutral theory of ecology supplies the appropriate null model for testing competition theory with respect to patterns SADs [species abundance distributions record the number of species and number of individuals per species], and (2), the neutral model fits a SAD well-enough and so fails to be rejected for that SAD, that therefore (3), either random drift, random immigration, or random speciation are the

dominant causes of that SAD.

I show that this line of argumentation fails to establish its conclusion because it (a), shifts from choosing between competition theory and neutral theory as exclusive alternatives to determining the unknown degree to which both are actually responsible, and (b), shifts from reasoning using the neutral theory in a way that is agnostic about the relation between neutral theory and a community to a way that assumes this relation is explanatory. In order to show this, in sections 3-6 I introduce a set of distinctions and with them construct a conceptual framework. In sections 7 and 8 I explain the context and theories involved in the line of argument, before showing that Hubbell made such shifts as I claim. In section 9 I finally apply the framework to the case of the neutral theory and show why the line of argumentation rests on a mistake.

2. Background to this chapter

(Note to the reader: This section is addressed to those with some knowledge of the use and discussion surrounding null and neutral models in biology. This section can be skipped with no loss to my argument. It can also be read after the conclusion.)

Both the kind of argument I am critiquing and my analysis of why it is mistaken have a history which I synthesize and build upon. Leigh van Valen identified my exact problem in a short scientific correspondence in *Nature*:⁷⁶

“In the past decade or so a subtle misuse of null hypotheses has become almost standard in ecology, biogeography, functional morphology and theoretical paleontology. ... The difficulty here comes when a null hypothesis is placed in a privileged position, to be accepted at least

provisionally until disproved.” (Van Valen 1985, 230)

The difficulty is that this privileging is illegitimate. I claim that this same difficulty is present in the use of the neutral theory of ecology. Yet van Valen offered no discussion of why it was made or why this is a mistake. My chapter furthers this discussion.

At the time, the Tallahassee Mafia (including most vocally Daniel Simberloff and Donald Strong) were trying to change the discipline of ecology. They pushed away from doing confirmatory science and toward doing good Popperian, falsificationist science by emphasizing the role of null hypothesis testing. Strong wrote,

“we propose another possibility with logical primacy over other hypotheses, that other hypotheses must first be tested against, but that is rarely considered at all by ecologists. This is the null hypothesis that community characteristics are apparently random” (Strong Jr, Szyska, and Simberloff 1979, 910)

Quinn and Dunham (Quinn and Dunham 1983) demonstrated that the way that the Tallahassee Mafia was arguing for using null hypotheses in a hypothesizing testing framework was mistaken. Rather than randomness having logical primacy for testing purposes, Quinn and Dunham showed that applied statisticians used null hypotheses rather to *estimate* the relative importance of factors *not* included in the null hypothesis, which are already *assumed* to be a cause.

Wimsatt (Wimsatt 1987) analyzed many ways false models (models which do not fit the ideal of perfect accuracy and unlimited scope, of which “neutral models” are a subset) are used to estimate the effect of processes not included in the false models. But while he agreed with how Quinn and Dunham viewed the role of testing null hypotheses, he described his position as:

⁷⁶ Thank you Chris Haufe for drawing my attention to this letter.

“Neutral models in biology represent “baseline models” or “null hypotheses” for testing the importance of efficacy of selection processes by trying to estimate what would happen in their absence.” (Wimsatt 1987 52)

In this Wimsatt was following the ambiguous terminology of the day. The volume *Neutral Models in Biology* (Nitecki and Hoffman 1987) in which Wimsatt’s paper appeared, opened with:

“Neutral model” belongs to a whole family of terms, which are sometimes, but by far not universally, used interchangeably; these are “null hypothesis,” “null model,” “random model,” “baseline model,” “stochastic approach,” “neutral theory,” etc. (3)

This is the morass Hubbell stepped into when he called his theory “neutral” and said it was “null”. The weakness of Quinn and Dunham’s critique today is that ecology has not followed their way of thinking about null hypotheses as baselines, but has rather followed the Tallahassee Mafia’s view that the problem context for using null hypotheses is hypothesis testing, not apportioning relative significance.

Beatty (Beatty 1987) showed how Kimura’s neutral theory of molecular evolution was being called a null hypothesis, and showed why it was no such thing. In evolution, the rhetoric of “testing null hypotheses” was used to support the neutralism research programme over its rival the adaptationism research programme. Beatty (Beatty 1997) then showed that in evolution the neutral theories and selection theories are often proposed as exclusive and alternative answers to the question of which theory to choose. But actually both neutral and selection theories were being used simultaneously to explain the same domains to apportion relative responsibility to random drift and selection.

Huss (Huss 2004) showed how the MBL in paleobiology, another neutral theory,

was taken to be both a null hypothesis and a null model. Gotelli and Graves (Gotelli and McGill 2006) addressed the differences between the neutral theory of ecology being *null* and being *neutral*. Odenbaugh (Odenbaugh) distinguishes between being *null* and being *causal*, and argues for the neutral theory being used as a causal theory in the way I characterize as *baseline modeling*. I base my distinctions between a null hypothesis, null model, and baseline model off of their discussions.

The present chapter extends this dialectic and builds upon Huss, Gotelli, and Odenbaugh in two ways. It better distinguishes the reasoning strategies followed when using the neutral theory as a null model vs using it as a baseline model. It also highlights the way in which two reasoning strategies are run together and across distinct problem contexts. Using this framework, I show how Van Valen's difficulties with privileging nulls are playing out in the ecology. And I show why it is mistake.

3. Theory Choice vs. Relative Significance

Consider two questions that scientists ask about a given domain:

Theory Choice Question: What general theory should we accept or pursue?

Relative Significance Question: What are the relative significances of multiple processes?

These questions are related but distinct. A Theory Choice Question asks for the correct or best theory of a given domain. One way of talking about underdetermination problems makes them about Theory Choice Questions. When empirical patterns alone do not decide, at least at present, between multiple sufficient explanations, theory choice is underdetermined by the data. Questions of this form assume that alternative theories are

mutually exclusive. Geocentrism and the Heliocentrism in astronomy and physics were rival answers to this type of question. A more nuanced form of a Theory Choice Question asks instead about pursuing a working hypothesis. On the latter version, a Theory Choice Question concerns the methodological choice of where to begin research, with no commitment to the superiority of the theory in the future.

A Relative Significance Question asks how to apportion relative responsibility among multiple, compatible processes. Here multiple processes are each assumed to be responsible to some degree for a given pattern. A simple case helps to distinguish a Relative Significance Question from a Theory Choice Question. Scientists are interested in the cause of the current warming trend of the Earth's temperature. One hypothesis is anthropogenic causes such as burning fossil fuels. Another is the earth's non-anthropogenic causes. Another is increased sun spot activity. But these are not exclusive alternatives and all or none of them could actually be responsible for our current warming trend. The research question is rather the relative significance of anthropogenic and non-anthropogenic causes.

John Beatty has argued that biologists primarily address Relative Significance Questions (Beatty 1997). As Gould and Lewontin said,

In natural history, all possible things happen sometimes; you generally do not support your favoured phenomenon by declaring rivals impossible in theory. Rather, you acknowledge the rival, but circumscribe its domain of action so narrowly that it cannot have any importance in the affairs of nature. Then, you often congratulate yourself for being such an ecumenical chap. (Gould and Lewontin 1979, 585)

Ecologists are interested in the relative significance of neutral processes compared to competitive processes in structuring particular communities. For example, Stephen

Hubbell has argued extensively that the tropical forest on Barro Colorado Island being is dominated by neutral processes. Ecologists are also interested in generalizations about the relative significance of neutral processes compared to competitive processes. For example, do tropical forest communities tend to be dominated by neutral processes? What about all tropical communities?

4. Null Modeling

The unified neutral theory of biodiversity and biogeography (UNT) asks the following question (Hubbell 2001): What community-level patterns result from the simplifying assumption that all species in a community exhibit the same demographic stochasticity (ecological drift) on a per capita basis, and ecological drift is the only process occurring besides random dispersal and speciation? ...

The value of posing this question rigorously is that we obtain a quantitative null hypothesis against which to test when, to what extent, and for which species demographic differences among species are necessary to explain observed community patterns. (Hubbell 2006, 1387)

Null modeling is a reasoning strategy used to test whether a process is causally relevant to a pattern. Null modeling takes place in the following context. A pattern calls for causal explanation. Scientists propose theory T that states that a set of processes P is responsible for the target pattern. They formalize the processes P into a mathematical model. They then establish whether the model is sufficient for producing the pattern. But even if sufficient, the underdetermination problem means that processes P need not be *actually responsible* for producing the pattern.

Null modeling is used to determine whether processes P is also *necessary* to produce the pattern. To do this, another model is constructed which formalizes processes N , where N is a strict subset of P . The model is *null with respect to* the processes A of the

alternative model. The prediction of the null model is then compared with the target data.

Once this is done, one of two inferences is made:

1. If the null model is sufficient to produce the target pattern, then
 1. the alternative model is rejected,
 2. the additional processes A are unnecessary to produce the target pattern, and
 3. the hypothesis that theory T correctly identifies the processes causally relevant to the target pattern is *undermined*.
2. If the null model is insufficient to produce the target pattern, then
 1. the null model is rejected,
 2. processes P are necessary to produce the target pattern, and
 3. the hypothesis that theory T correctly identifies the processes causally relevant to the target pattern gains positive evidence.

Null modeling is *not* a method for choosing between two theories, but a method for evaluating one theory. The two models are treated asymmetrically. The alternative model is tested by the null model, but not vice versa. Only the model playing the role of the alternative model is in a position to gain positive evidence. However, there is nothing preventing a model from in one case playing the role of the null model and in another case playing the role of the alternative model. Models are not inherently null, although particular models are inherently appropriate or inappropriate for playing the role of a null with respect to another model.

A simple example makes this characterization concrete. Attempting to explain the species-to-genus ratio (S/G) brought *null modeling* to ecology.⁷⁷ Darwin thought that the struggle for existence would be strongest between members of the same genus because they have similarity in habits, constitution, and structure. This led biologists to expect the S/G to shrink as resources became fewer, since fewer species within a genus would be able to sustain themselves. One extreme case occurs when competition is so strong that $S/G = 1$, and one species excludes all others in its genus. Elton (Elton 1946) analyzed S/Gs in published surveys of plant and animal communities from a variety of habitats. He found that the average S/G in small communities, such as islands, were considerably smaller than S/G for the larger region. This pattern called for explanation. Elton proposed that differences in S/Gs were best explained by differential competition between species of the same genus (*intrageneric competition*).

Williams (Williams 1947, 1964) undermined Elton's explanation using null modeling. Williams claimed that ecologists had so far merely *assumed* that a smaller S/G on an island than in its surrounding area was evidence for selection against congeneric species. He argued,

“This is, however, a statistical error as, with a truly random sample of a small number of species from a larger population, there is always a reduction in the average number of species per genus.”

(Williams 1947, 256)

Williams demonstrated this error by conducting a simulation experiment with a model.

⁷⁷ This case is used to introduce null modeling in (Gotelli and Graves 1996) and (Huss 2004).

But this is the second round of the null model war. The S/G was debated in much the same way between Jaccard and Palmgren in the 1920s in French and German journals. See (Jarvinen 1982) and (Cadotte 2015).

Williams compared the data from the observed number of species of flowering plants of Britain in larger and smaller areas. The total number of species of flowering plants in Britain was known to be approximately 1251, and so Williams placed 1251 discs in a box and, ‘after prolonged shaking’, drew samples of different sizes and recorded the number of genera.⁷⁸ He did this three times for each sample size and averaged the results.⁷⁹

Williams showed that the number of genera in a sample produced by a process *random with respect to genera* was actually greater than found in an observed sample of the same size.

Elton had assumed that intrageneric competition would increase the S/G in an island community, since islands tend to have more homogeneous conditions than the surrounding area and species of the same genera tend to utilize similar resources.

Williams argued that, because a pattern like the one Elton observed could be achieved with a model null with respect to generic relations, generic relations were unnecessary.

Williams did not further argue that this was positive evidence for the causal *irresponsibility* of generic relations to the S/G. This case follows the way I characterized the inference pattern used when null modeling, when the null model cannot be rejected.

Before moving on, two clarifications need to be made. First, there is an important difference between *null modeling* and statistical *null hypothesis testing*. The relevant difference for this chapter is that null modeling is used to test whether a process is

⁷⁸ Williams doesn’t say, but I imagine that the discs were of different colors representing genera.

⁷⁹ Huss (Huss 2004, 124) points out that this is a non-computerized version of a Monte Carlo experiment.

causally relevant to a pattern, while null hypothesis testing is used to determine whether there is a pattern to be explained. Another way to express this is that neutral theory is used as a *substantive theory*, not a *statistical theory* (Mayo and Spanos 2006).

Second, null modeling often relies on *parameter-fitting* the null and alternative models, not independent parameter estimation. Null and alternative models are then assessed on their capability to fit the data. A null model is accepted or rejected based on whether its best-fit is good enough to meet the standards set. Even though not any model can fit any pattern arbitrarily well, this obviously biases the evaluation procedure in favor of accepting the null model when it does not accurately fit the target system. But this problem is different from the main problem I address in this chapter, though it is related and compounds it, as I explain in section 8.

5. Baseline Modeling

“Neutral models provide a baseline expectation for important questions, such as the effect on biodiversity of habitat fragmentation.” ...

“Neutral theory provides a valuable null model, a baseline to compare with empirical data and other models. If a neutral model is inconsistent with data, one can learn something about what missing biological factors are needed to improve it.” (Rosindell, Hubbell, and Etienne 2011, 7)

Baseline modeling is a reasoning strategy used to apportion relative causal responsibility. Baseline modeling takes place in the following context. Scientists accept that multiple processes are jointly responsible for producing a (set of) pattern(s). They want to determine the relative contribution of each process. Baseline modeling proceeds by proposing that one (set of) process is responsible for the pattern in the absence of the other processes. This *baseline state* is formalized by the *baseline model*, which gives the

baseline expectation. Deviations from the baseline expectation are explained by invoking *additional* processes. This estimates the relative significance of baseline and additional processes to the pattern of interest.

When baseline modeling, you measure the baseline responsibility first, and use the additional processes to fill in the gap. In this way, baseline modeling proceeds by *assuming* that the baseline model accurately describes the system, both in the complete absence of additional processes, and to a degree inversely proportional to the strength of additional processes. Models are not inherently baselines, although particular models are inherently appropriate or inappropriate for playing the role of a baseline with respect to certain processes.

Again, a simple example makes this characterization concrete. Newton's first law, the principle of inertia, is a quintessential baseline. The principle of inertia can be stated as: a body will maintain uniform speed in a straight line unless acted upon by an outside force. Using this law with definitions, one can infer the presence of forces from the fact of non-inertial motion.

A basic problem in classical mechanics is to determine force attracting a body of some mass orbiting in uniform circular motion at a known radius around a much more massive body. One solves this problem by using the formula for Newton's law of gravitation. A background assumption for the problem is that the orbiting body would have perfect inertial motion if the massive body were not present. In this way inertial motion is used as the baseline expectation in a causal explanation invoking forces.

Baseline models can also be used hierarchically to solve problems. Additional

force laws can be added to the principle of inertia. To detect further forces at work in the system, one then looks at deviations from what one expects given both inertia and the additional forces already accounted for. The discovery of Neptune was aided by the calculations by Urbain Le Verrier and John Couch Adams in this way. The motion of Uranus deviated from the expectation given only the known planets and they each calculated approximately where the new planet should be.

6. Interrelations in the framework

With the four categories of my two distinctions explained, it is now important to get clear on the six ways in which they are related to each other. But first, a note on my conceptualization. I am using the concepts *model* and *theory* in the following way.

Theories have a causal interpretation and are compared indirectly to data via a model.

Models gain their causal interpretation through their connection to the theory they formalize. Therefore, theories are required for causal explanation, but models will also be required depending on the reasoning strategy being used. The relationship between a theory and a model that formalizes it is typically many-many. There will be multiple ways of formalizing a set of causal processes into mathematics. And there will be multiple causal interpretations about a set of equations. Nevertheless, null and baseline modeling both rely upon two basic inferences: from the similarity between the output of a model and the data, infer the similarity between the model and the target system. From this, infer that the theory supplying the model identifies the processes causally relevant to the target pattern in the target system (Giere 1998).

The six relations are the following:

(1) *Theory Choice Question and Null Modeling.* A Theory Choice Question asks which general theory we should accept or pursue. This is an exclusive choice between multiple proposed theories. Null modeling is therefore not suited to answering a Theory Choice Question. Null modeling is a strategy for reasoning about whether *one* theory, the theory that supplies the alternative model, identifies processes causally relevant to the target pattern.

Null modeling is a method for answering whether the theory supplying the alternative model has positive evidence. While null modeling *seems* to be appropriate for deciding between two theories via nominating their respective models as *alternative model* and *null model*, this is not so. Only the theory which supplies the alternative model is in a position to be evidenced and thus to be accepted or pursued by a Theory Choice Question.

(2) *Relative Significance Question and Baseline Modeling.* A Relative Significance Question asks how to apportion relative responsibility to multiple processes contributing to one pattern. Baseline modeling is perfectly suited to answering a Relative Significance Question. The main difficulty here is how to decide which set of processes generate the baseline expectation. Given a baseline expectation, deviations can be detected. Then the ratio of the responsibility of the baseline processes to the additional processes is the ratio of the degree of agreement with the baseline model to the degree of disagreement with the baseline model. Deviations are explained by invoking additional processes *and* the baseline processes.

Baseline modeling is not the only method for answering a Relative Significance Question. There is no reason why apportioning causal responsibility must be based on

detecting deviations from baseline processes. Some think that there is no principled way to decide which processes should function as the baseline (see for example, (Lloyd 2015) on an analogous issue in evolution).

(3) *Theory Choice Question* and *Relative Significance Question*. A Theory Choice Question is epistemically prior to a Relative Significance Question because evidence is prior to explanation.

(4) *Null Modeling* and *Baseline Modeling*. The formal similarity of null and baseline modeling leads to their conflation. Both reasoning strategies use two models and treat them asymmetrically. The null model and baseline model are both compared with the data prior to their counterparts, and their degree of fit is used to inform a judgment about the alternative model and additional processes. The important difference is that null modeling *is agnostic about* the relationship between theory supplying the null model and the target system, while baseline modeling *assumes* that the relationship between the baseline theory and the target system is explanatory.

(5) *Theory Choice Question* and *Baseline Modeling*. Accepted or pursued theories can be used to give causal explanations. Baseline modeling is one way of giving causal explanations. But additional judgments must be made, most importantly, about which processes are the baseline and which are additional. This choice is difficult and there is no agreement about whether it is a purely *empirical* choice or not. For example, was the switch from Aristotle's view that rest was the natural state of heavy matter to Newton's view that inertial motion is the natural state a matter of only better observation and

experiment? This is not obvious, but it will not be pursued in this chapter.⁸⁰

(6) *Relative Significance Question* and *Null Modeling*. Null modeling can be used to inform whether *one* theory is evidenced or undermined. An evidenced theory in turn can be used to help answer a Relative Significance Question. But additional judgments are required. Null modeling is in general insufficient to eliminate all but one theory from a set. Also, answering a Relative Significance Question does not require baseline modeling.

7. The neutral theory of ecology

Two of the most extensively studied patterns in community ecology are the relative species abundance distribution (SAD) and the species-area curve. In this chapter I focus on SADs. A SAD records both the number of species and the number of individuals within each species in a single-trophic level community. A single trophic-level community is all the organisms which are at the same energy level in the same place. All the trees and shrubs in a forest are an example, as are the nesting birds on an island, if the raptors are excluded. A SAD is presented as a curve. Figure 4.1 compares five SADs (from (Hubbell 2001)).

⁸⁰ Experiments by Hooke showed that the time a top can spin is proportional to its weight and inversely proportional to the resistance with the ground and the air. This suggested that, in the absence of resistance, the top would never stop spinning. But D'Alembert argued that the first law was actually known a priori and followed from the definitions and axioms. Whewell argued that, while it almost followed only the axioms of cause and definition of force, empirical data was required to establish that time itself would not diminish the velocity of the top. See (Whewell 1847, 217).

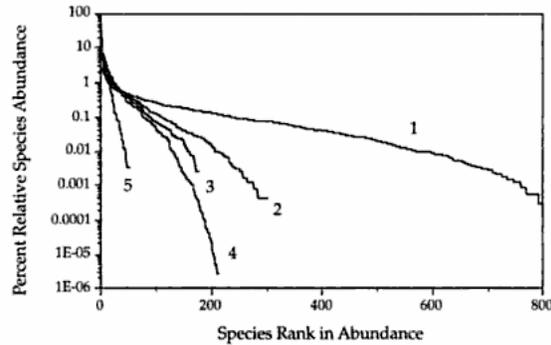


FIG. 1.1. Patterns of relative species abundance in a diverse array of ecological communities. Species in each community are ranked in percentage relative abundance from commonest (*left*) to rarest (*right*). The percentage relative abundance is log transformed on the y-axis. 1: Tropical wet forest in Amazonia. 2: Tropical dry deciduous forest in Costa Rica. 3: Marine planktonic copepod community from the North Pacific gyre. 4: Terrestrial breeding birds of Britain. 5: Tropical bat community from Panama.

Figure 4.1 – Comparison of species abundance distributions
From (Hubbell 2001, 4). Used with permission from Princeton University Press.

A single point on a line is read as: the x th ranked species makes up $\log_{10}(y)\%$ of the community. These curves have an ‘S’ shape - steep, then flatter, then steeper. This reveals how most species are relatively rare and few are common. Ecologist McGill (McGill et al. 2007, 996) has called this one of community ecology’s few descriptive laws. This pattern is robust across many different communities and so calls for a common causal explanation. Both competition theory and neutral theory are used to give such causal explanations.

Most ecologists working on these problems accept that the most important causal factors relevant to these patterns are the competitive differences between individuals of different species. The basic idea begins with thinking that every species is best adapted to a particular set of resources and conditions, called its *niche*, from natural selection. When two species live in the same area, the species better at utilizing the limited available

resources will tend to out-compete the other. This idea has been generalized to form the *competitive exclusion principle*, competition theory's basic principle. It states that if there are insufficient differences between how two species' utilize resources and effect their environment, then those two species cannot stably coexist. One species will come to exclude the other. Philosopher Weber (Weber 1999) has been called this principle a causal law. The general process responsible for this is *interspecific competition*, which is seen both as the ecological analog of natural selection and as a result of natural selection. This general approach in community ecology is called competition theory, and it has proven very successful.⁸¹

As I show in the next section, Hubbell explicitly constructed neutral theory to counter proposals that explained patterns such as SADs by invoking interspecific competition. If the neutral theory accurately describes the significant causes of SADs, then competitive exclusion is not a strong cause of SADs. This follows from the character of neutral theory, and from its use as a baseline model.

The *neutral theory of ecology* has two defining characteristics. First, it posits drift (random birth and death), immigration, and speciation. Collectively called *dispersal processes*, these control how current SADs can change over time. Dispersal processes are chancy, historical processes because the SAD at a time is the result of both the SAD at a previous time and the probabilities of an individual dying, giving birth, immigrating, and speciating.

⁸¹ Two examples of competition theories are Tilman's R* Theory (Tilman 1980, 1982) and Chesson's Coexistence Theory (Chesson 1994, 2000).

Second, it assumes *neutrality*. Neutrality states that all individuals in a community, regardless of species, have identical ecological rates. In a neutral-dispersal theory individuals have identical chances of dying, giving birth, immigrating, and speciating. By assuming neutrality, the process of interspecific competition is ruled out, since this requires competitive differences between species.

A *neutral model* of the neutral theory of ecology is a mathematical model which formalizes the assumptions made and processes discussed by the theory. The processes described by the basic neutral model (Hubbell 2001), sometimes called the island-mainland model, can be visualized with the figure 4.2 (from (Rosindell, Hubbell, and Etienne 2011)):

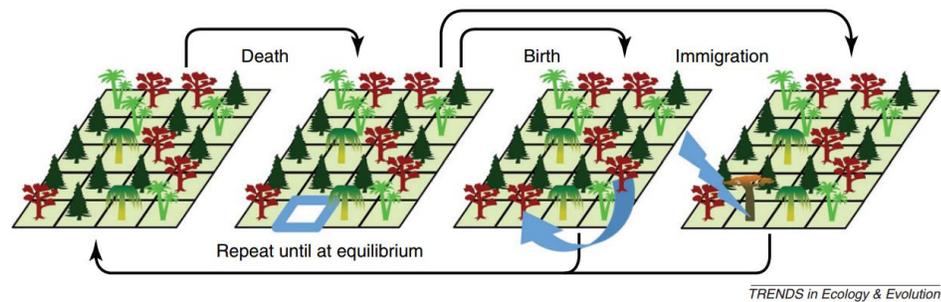


Figure 4.2 – Death, Birth, Immigration cycle in Neutral Theory
From (Rosindell, Hubbell, and Etienne 2011, 2). Used with permission from Elsevier.

The mathematics of neutral model are interpreted as follows:⁸² imagine a

⁸² Models of this type have a history in ecology. They are formally similar to models from population genetics, especially developed by Kimura for his Neutral Theory of Molecular Evolution (Kimura 1983). These models were reinterpreted ecologically and imported into community ecology by Caswell (Caswell 1976) and Hubbell (Hubbell 1979). But they also stem from modeling developed to expand MacArthur and Wilson’s Theory of Island Biogeography (MacArthur and Wilson 1967), especially by the Woods Hole Group’s MBL Model of clade diversity, which balanced speciation and extinction rates (Raup et al. 1973). They were then developed much further in the infamous Null Model Wars between Diamond (Diamond 1975) and Simberloff (Connor and Simberloff 1979, 1132). For a more complete history focusing on the MBL model, see Huss (Huss 2004). For a more complete history of the neutral theory of ecology, see Chs. 1 and 2 of this dissertation.

community as a checkerboard, each space represents some place for an individual to live. Fill the board with pieces of various colors drawn from a bag. Each piece represents an individual and each color represents a species. The board represents the island or local community and the bag represents the mainland or metacommunity. The neutral theory models how the SAD changes over time and predicts the equilibrium SAD. Each turn consists of removing a piece and replacing the space with a new piece to maintain constant community size. To begin, remove a piece at random (i.e. with equal probability for each piece), this represents a death. The death can either be replaced by a local birth or by immigration. With probability $1-m$, replace the death with a piece according to the following rule: each piece currently on the board has an equal chance of reproducing and filling the space. Therefore, the probability that a given color will fill the space is proportional to the current relative abundance of that color. This follows from assuming *neutrality*, that all individuals in the community are functionally equivalent. With probability m , replace a death according to the following rule: each piece currently in the bag has an equal chance of immigrating and filling the space. For this the dynamics of a metacommunity must also be tracked, and these follow analogous rules. Here drift occurs and is balanced by speciation, instead of immigration, where a death in the bag is replaced by an individual of a novel species. After many rounds of death and replacement, an equilibrium (time-independent) SAD will be reached.

The basic model has three free, interpreted parameters: J , local community size; m , immigration rate; and θ , a function of J_M , metacommunity size, and v , speciation rate. The model predicts that the equilibrium patterns will be statistically indistinguishable from a particular statistical distribution.

Figure 4.3 shows an example of a very close fit between observed SADs and the output of the neutral model:

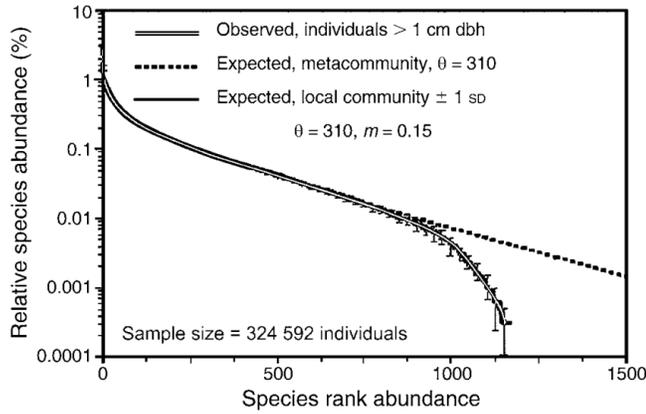


FIG. 1. The fit of the unified neutral theory of biodiversity and biogeography (UNT) to the dominance–diversity curve for the tropical tree community in Lambir Hills National Park, Sarawak, Borneo. The dotted line extending diagonally down to the right is the best-fit metacommunity curve for $\theta = 310$, assuming no dispersal limitation (probability of immigration $m = 1$). The parameter θ is the fundamental biodiversity number of neutral theory, which is a composite of the speciation rate and the size of the metacommunity source area. The distribution of relative tree species abundance for the 52-ha plot was best fit with $\theta = 310$ and $m = 0.15$. The heavy solid line is the observed dominance–diversity curve. The agreement between the fitted line and the observed line for 1197 species is excellent ($r^2 = 0.996$). Error bars represent \pm SE. Note the log scale of the y-axis.

Figure 4.3 – Close fit of Neutral Model to BCI SAD
From (Hubbell 2006). Used with permission from John Wiley and Sons Inc.

The double solid line is the observed SAD of a tropical tree community in Lambir Hills Park, Sarawak, Borneo. Two expectations from the neutral theory are compared. In both cases, local community size J is known, but the values of the other free parameters are those that produce the best fit to the SAD (mentioned at the end of Section 3). The dotted line shows the dispersal-unlimited version of the neutral model, where the immigration rate m is fixed at 1. The single solid line shows the dispersal-limited version of the neutral model, where immigration rate is also a free parameter. The dispersal-limited model fits very well, with a $r^2=0.996$. Besides producing such close fits using independent estimates of the values of free parameters for a community, this is the best possible outcome for the neutral theory. Given close fits such as this, what conclusions should we infer? The answer to this depends upon the reasoning strategy being used. In the next section I show how Hubbell shifted his reasoning strategy from null modeling to baseline modeling, and how the conclusions he inferred from close fits to SADs also

changed.

8. The use of neutral theory

In this section I show that Hubbell shifted from using the neutral model as a *null model* for testing competition theory with respect to SADs, to using the neutral model as a *baseline model* for measuring the strength of neutral-dispersal and competitive processes with respect to SADs.

Hubbell deployed the ancestor of neutral theory, the *community drift model*, to undermine the hypothesis that SADs of tropical rain forests are dominated by interspecific competition (Hubbell 1979, Hubbell and Foster 1986). Hubbell drew two conclusions from the comparison of the community drift model with the SAD of a tropical forest plot in Costa Rica.

“First, we may expect to observe substantial differentiation of the relative abundance of species in natural communities as a result of ... a kind of “community drift” phenomenon.” (Hubbell 1979, 1307)

By varying the free parameters of the model, a wide variety of SADs can be approximated. This established that the community drift model can fit some observed SADs. It is surprising that this variety can be achieved by random drift and immigration because it conflicts with expectations conditioned by explaining SADs with interspecific competition.

“Second, we cannot necessarily conclude that, just because a species is of rank-1 importance in a community, its current success is due to competitive dominance ... stemming from some superior

adaptation to the local environment.” (Hubbell 1979, 1307)

The ability of the community drift model to fit to observed SADs shows interspecific competition is not required to explain either SADs or the fact that a particular species is most dominant. So far, the community drift model was only used as a null model to undermine support for interspecific competition being responsible for SADs.

In 1997 and 2001, Hubbell presented neutral theory, now incorporating speciation, as a general, positive *theory* of abundance and diversity (Hubbell 1997, 2001). Hubbell began to use neutral theory as a baseline model with respect to interspecific competition, but he did not distinguish this new usage from its use as a null model. This can be seen in the following exchange.

Terborgh et al. used the community drift model as a null model for testing interspecific competition (Terborgh, Foster, and Nuñez 1996, 564). They observed that particular species tended to be the commonest species across a set of five forest plots. Because they thought that this was very unlikely given the community drift model, they reasoned that this correlation established that these species were better competitors in that region.

Hubbell responded that, if the five sites are connected by immigration (which was reasonable because they all occurred in the same river valley), neutral theory predicts that the most abundant species in one local community will be the commonest species in another local community “essentially 100% of the time” (Hubbell 2001, 335). Terborgh et al. incorrectly thought that there should be no spatial correlation in abundance on the

community drift model. Therefore, the community drift model qua null model should not be rejected and competitive differences should not be invoked. So far, this is consistent with neutral theory supplying only a null model.

Hubbell further argued that, while some correlation between the commonest species did not show the importance of interspecific competition, the *degree of* correlation in which the commonest species dominate the other species in abundance did show this. By comparing the observed correlation with the expectation given neutrality, Hubbell estimated that this ‘ecological dominance deviation’ showed a 6% fitness advantage in the commonest species (Hubbell 2001, 337). Here Hubbell used neutral theory as a baseline model.

It is worth noting, in service of making clear how intertwined talk of null hypothesis testing, null modeling, and baseline modeling are, that when Hubbell made this response, he wrote,

“What they did not comment on, however, was the excessive dominance itself—presumably because they had no prior statistical hypothesis of what null relative abundance distribution to expect.” (Hubbell 2001, 336)

This comment makes clear the value of dominance detection, made possible by baseline modeling.

Most community ecologists agree that the Theory Choice Question has been answered. All four processes - drift, immigration, speciation, and competition - are causally relevant to the patterns of interest.⁸³ The open question is a Relative Significance Question. Hubbell’s explanatory strategy for tackling this question uses neutral theory to

detect deviations and assign responsibility for the agreement to neutral theory. But his justificatory strategy does not distinguish between null modeling and baseline modeling. And this is a problem because, as I show in the next section, it illegitimately privileges neutral theory as accepted until rejected.

9. Persuaded of these principles, what havoc must we make?

Examining the history of neutral theory in light of the distinction between null modeling and baseline modeling shows that *neutrality* was used to shift answering a Theory Choice Question to answering a Relative Significance Question. But the two questions must be kept distinct, and neither null modeling nor baseline modeling is alone sufficient to address both questions.

The target of my critique is the line of argumentation that privileges neutral theory in the following way. Begin with the Theory Choice Question of competition theory vs neutral theory. Because neutral theory ignores species differences, it is used to test competition theory. The *burden of disproof* is on competition theory to disprove neutral theory before competition theory can be accepted. Further, the *benefit of the doubt* is given to neutral theory in the sense that, if neutral theory is not rejected, then it should be accepted and used to give causal explanations. If neutral theory proves adequate for describing some SADs, then causal conclusions of the following form are drawn:

“Species interactions, niche partitioning, or density-dependence, while they may be present, do not appear to enhance tree species richness at Barro Colorado.” (Condit, Chisholm, and Hubbell

⁸³ See (Vellend 2010) for the theoretical development of this ‘conceptual synthesis’ framework. Vellend has a book coming out soon that extends this framework of the four processes.

2012a, 1)

This is a negative conclusion about the lack of responsibility of several processes traditionally associated with competition theory in the tropical rainforest on Barro Colorado Island. Positive conclusions are further drawn,

“The neutral model predicts diversity and abundance at Barro Colorado because it properly describes what matters most – species input [neutral speciation] – while ignoring irrelevant details.” (Condit, Chisholm, and Hubbell 2012a, 5)

The implied justification of burden of disproof and benefit of the doubt are the *neutrality* of neutral theory and its ability to provide the appropriate null model for testing competition theory. As Hubbell said elsewhere, by using neutral theory,

“we obtain a quantitative null hypothesis against which to test when, to what extent, and for which species demographic differences among species are necessary to explain observed community patterns” (Hubbell 2006, 1387)

It is useful to reconstruct the line argumentation in the following way for analysis:⁸⁴

Argument 1

1. The neutral theory of ecology supplies the appropriate null model for testing competition theory with respect to SADs.
2. The neutral model fits a SAD well-enough and so fails to be rejected for that SAD.
3. Therefore, either random drift, random immigration, or random speciation are the dominant causes of that SAD.

⁸⁴ This is my rational reconstruction. No one has stated it so cleanly.

Assume we are dealing with a case where premise 2 is true. Line 3 is supposed to follow from 1 and 2. Because the neutral model qua null model is not rejected, we are supposed to infer that neutral processes dominate. Because null modeling is unsuited to answering a Theory Choice Question, this argument is off to a bad start. But this is only one problem with the line of argumentation.

Argument 1 also shifts from a Theory Choice Question to a Relative Significance Question. While premise 1 is about testing or validating competition theory, the conclusion, line 3, shifts to answering the relative significance of neutral to competition processes. The shift is obvious if we ask how Argument 1 would proceed were premise 2 false. If we are addressing a Theory Choice Question, then it seems we should infer that competition theory is just not rejected. If we are addressing a Relative Significance Question, then it seems we should infer that interspecific competition dominates and neutral processes are relatively unimportant. In the case where neutral theory is empirically adequate, the steps are run together.

The following would be an extreme form of running the two questions together, compounded by the issue of parameter-fitting mentioned in Section 3⁸⁵. Competition theory and neutral theory are tested against each other by determining which theory is

⁸⁵ This is done by (Chen 2016). Chen compares the basic neutral model with a model incorporating local speciation, which adds one free parameter. Using AIC, the local speciation model gets a slightly higher AIC score.

Chen then moves from testing the two models using BCI FDP data, to using the best-fit parameter estimates as a *measure* of the relative strength of speciation vs local speciation on the BCI FDP:

“Therefore, local speciation rate v_{local} was at least 10-fold larger than the background speciation rate $v_{\text{background}}$.” (Chen 2016, 69)

capable of better best-fits of a set of SADs. The best-fit between a model and SAD is determined by parameter-fitting the models to yield the closest fit between model and data. Before the testing, some standards are set for how much better the best-fit of competition theory than neutral theory in order to reject neutral theory and accept competition theory. Imagine that those standards are not met and neutral theory is not rejected. Because the neutral theory functioned as a null model for testing competition theory, this gives no confirming evidence to neutral theory. But it is often mistaken as confirming evidence and said to mean neutral theory is accepted.

Now the same estimates of parameter values which yield the best-fit between neutral model and a given SAD are taken to be a measurement estimate of the correct values of immigration and speciation rate. And because of the flexibility of the neutral model, the best-fit to a given SAD is probably pretty good. This mistakenly supports the conclusion that drift, immigration or speciation are the dominant causes of abundance and diversity in the community. This is clearly bad reasoning.

One source of the error is treating the same statistical procedure for evaluating whether a model should be accepted or rejected as is used to measure parameter values. There is no good reason and some bad reasons to do this.⁸⁶ Hence this is an extreme form of running together a Theory Choice Question with a Relative Significance Question. A second source of error was in taking the failure of the neutral model to be rejected as anything more than a disconfirmation of the competition model. The correct attitude to take after failing to reject the null model in a test is agnosticism towards the null.

⁸⁶ See (Mayo and Spanos 2006) on why the probabilities in each mean different things. They argue that NP doesn't use confirmation, but error-probabilities.

If ecologists want to gain evidence for the neutral theory using null modeling, then the neutral model must play the role of the alternative model and be tested by an appropriate, more minimal null model.⁸⁷ If neutral theory passes the test, it gains evidence (or stays zero if you are Popperian). If neutral theory fails the test, it loses some (or possibly all) of its evidence. One simple way to do this would be to test the neutral theory against the community drift model with respect to SADs. Finding the community drift model insufficient to fit a SAD, would provide evidence for speciation.

Therefore, the stalemate between niche and neutral theory cannot be broken by only using the neutral theory qua null model to test competition theory. Null modeling can inform a Theory Choice Question only by iterating the process while varying the theory being tested and the theory supplying the null model. But in the process, we shift from comparing competition theory with neutral theory to neutral theory with a new, more minimal theory.

Finally, the reasoning used to infer causal conclusions, as reconstructed in Argument 1, on the basis of using the neutral model qua null model is invalid. In Argument 1, line 3 does not follow from lines 1 and 2 because failing to reject neutral theory qua null model only says something about the support for competition theory. Also, in moving from line 1 to line 3, we switched from a Theory Choice Question to a Relative Significance Question illicitly. If causal conclusions about the strength of neutral and competitive processes should be drawn, then they must follow from different lines of argument.

The most charitable interpretation of Argument 1 and the associated reasoning is

⁸⁷ Gotelli and McGill recommend this procedure. (Gotelli and McGill 2006)

the following. Despite the language used, baseline modeling is actually being used to answer a Relative Significance Question. This is perfectly acceptable, though the inferences drawn would be stronger if independent estimates of the parameter values were available. Therefore, the open question is what reasoning supports assuming that neutral processes are actually a cause of the patterns of diversity. There is no reason this must come from null modeling.

10. Conclusion

I have shown that a subtle mistake is being made in the use of the neutral theory of ecology that undermines the causal conclusions being drawn. There is an almost irresistible allure to using a ‘neutral’ theory as a ‘null’ and then failing to distinguish null hypothesis testing from null modeling from baseline modeling. You cannot use null modeling to answer a question about the relative responsibility of multiple causal factors. As with other fallacies, or a magic trick, once you see the trick in action, it is impossible to miss it. I do not challenge the *truth* of the causal conclusions drawn, which are supported by many different lines of argumentation. I challenge only validity of the line of argumentation which shifts from using neutral theory as null model to baseline model and shifts from answering a Theory Choice Question to a Relative Significance Question.

Chapter 5.

False Models for Fecund Research Programmes

1. Introduction

This chapter illuminates three aspects of the neutral theory of ecology using concept of the research programme: (1) the synchronic uses of the neutral theory to make predictions and give descriptions and explanations; (2) its diachronic development in response to theoretical innovation and confrontation with data; (3) the fault lines of its complex relationship to competition theory.

Previous chapters have focused on the synchronic use of the neutral theory. I have focused in particular on the ways that its proponents use it as a ‘null hypothesis’ for testing competition theory in order to give neutral theory the benefit of the doubt and competition theory the burden of disproof. I showed that while the neutral theory is fruitfully used as a *null model* with respect to competition theory, this use is of no use for inferring causal conclusions about the relative strength of neutral-dispersal and competitive processes. These causal conclusions instead follow from using the neutral theory as a *baseline model* from which deviations from the predictions of the neutral model can be detected. This chapter shows this synchronic use of neutral theory for *baseline modeling* is something done by the neutralist research programme as part of its diachronic development.

Foregrounding the use of the neutral theory as a baseline model had the effect of backgrounding the diachronic development of the neutral theory. This chapter inverts that

emphasis in order to show baseline modeling as one heuristic among several for responding to confrontations of the neutral theory with data. The neutral theory accurately describes some patterns for some communities. These are clear successes. But when the neutral theory inaccurately describes a pattern, neutralists do not respond rejecting the theory. They rather respond in one of two ways: either they take this as a sign that competitive processes are causally important to the pattern (using the Baseline Modeling heuristic), or else they take it as a sign that there is a problem with the neutral theory or its formalization (which I call the *Adding Complexity* heuristic). The Adding Complexity heuristic is a main driver of the diachronic development of the neutral theory. Adding Complexity works in concert with Baseline Modeling, and both are manifestations of a methodological commitment of the research programme to approach understanding ecological communities from the starting hypothesis that only neutral demographic and dispersal processes are causally responsible for observed patterns of abundance and diversity.

Finally, the neutral theory was developed as an alternative to competition theory. Competition theory is well-developed and successful. Ecologists have argued since Hubbell published his 2001 book about what any successes of the neutral theory say about the standing of competition theory. The relationship between the theories is complex in part because both continue to be developed. Some ecologists have characterized their relationship as an exclusive either-or theory choice question to be decided by a crucial experiment refuting neutral theory (for example (Dornelas, Connolly, and Hughes 2006)). Other ecologists have focused on integrating the theories in order to solve a relative significance question (Vellend 2010). For the latter ecologists,

the synthetic theory hypothesizing the processes of drift, immigration, speciation, *and* competition is clearly best situated to answer the actual relative significance questions being addressed.

But it is too quick to claim that a conceptual synthesis of competition and neutral theory alone will eliminate the controversy. One virtue of the former, exclusivist view is that competitionists and neutralists have distinct and incompatible methodological commitments. However, this was not the focus of the exclusivists because the exclusivists mistakenly focused on the *theories* in play from a Popperian view of science as conjecture and refutation. This led the exclusivists to seek crucial experiments that could reject one of the theories. But the methodological disagreement is not in the theories, but in their respective research programmes. Both competitionists and neutralists accept that drift, immigration, speciation, and competition are all real causes and that each is sometimes the dominant cause of a pattern. What they disagree about is the methodological question of how to go about determining the relative significances. Competitionists look for competition first; neutralists look for neutral-dispersal first. And so a methodological question faces the users of any synthetic theory that incorporates all the accepted factors - which processes should you look for first?

This is the argument of this chapter in brief. Section 2 begins with a brief historical introduction to competition theory and neutral theory. Section 3 presents a history of 20th century philosophy of science to argue that ecologists currently inhabiting a Popperian view of science would be better suited as Kuhnians or Lakatosians. Section 4 draws on a distinction between three kinds of adaptationism present in evolutionary

biology in order to distinguish between three kinds of competitionism and adaptationism. Section 5 characterizes the neutralist research programme. Section 6 argues that the primary disagreement between neutralists and competitionists is not empirical but methodological.

2. A brief history of competition theory and neutral theory in ecology

One general question addressed in community ecology is how to resolve the *biodiversity paradox* (Hutchinson 1961). The paradox arises from the tension between two seemingly true, but incompatible, claims. The first is a theoretical principle. The second is an empirical generalization.

Competitive exclusion: Similar species cannot stably coexist, because one will come to exclude the other through competition.

High Species Richness: In many communities, similar species coexist.

Their incompatibility is clear. If the Competitive Exclusion Principle is true, then there should not be high species richness for ecologically-similar species. But since we observe high species richness for similar species, therefore the similar species must be able to coexist. Similarly, if High Species Richness is true, then similar species can coexist. But since the competitive exclusion principle is an established principle, therefore species richness must not be high for similar species. Ecologists have developed competition theory and neutral theory in order to provide, among many other entangled issues in ecology, resolutions of this paradox.

In the following simplified history, I first describe how the paradox arose, and

describe how the competitionist and neutralist resolutions to the paradox began. This historical development in particular some of the distinct goals and decisions made by the two groups in the face of understanding some of the most complex systems in nature.

Competition has been hypothesized to be an important ecological process for a long time. Looking for the effects of competition led to the development of the competitive exclusion principle. Darwin suspected that intrageneric competitive exclusion took place (Darwin 1859). Tansley and Gause did early experimental work that showed one species would out compete another for a limited resource (Tansley 1917, Gause 1932, Gause 1934). Grinnell introduced the concept of the *niche* to explain the geographical distributions of species. Grinnell's niche concept encompassed both the habitat type a species lived in, and the physical and behavioral adaptations of the species that enables it to live (Grinnell 1917b, a). He first stated what came to be known as *the competitive exclusion principle*, "It is, of course, axiomatic that no two species regularly established in a single fauna have precisely the same niche relationships" (Grinnell 1917b, 433). Elton defined the niche rather as the place of a species in its biotic environment, including resources and competitors, and its effect on the environment (Elton 1946, Elton 1927). He also used the idea of competition between niches to explain the abundance and distributions of species. These two senses of 'niche' were merged together to include the abiotic and biotic conditions of existence, and the physical and behavioral traits, of a species. This is the sense of 'niche' meant when the competitive exclusion is now expressed as if two species share the same niche, then they cannot stably coexist.

The general competitionist approach to resolving the biodiversity paradox is to deny the empirical claim that similar species have high species richness. One strategy looks for important differences in how species previously seen as ecologically similar actually utilize resources and effect the environment. Another strategy looks for previously unappreciated heterogeneities in the distribution of resources in space and time which different species may differently track (Tilman and Kareiva 1997). These two strategies are compatible. Either way, if species 1 is better able to utilize resource R than species 2, then species 1 should be more abundant than species 2 when R is limited. In what follows I speak in terms of the first strategy.

Lotka and Volterra each did early foundational work and developed simple mathematical models of how two species compete with each other for a resource and when one will exclude the other (Lotka 1925, Volterra 1926). Robert MacArthur extended this work in one influential direction when he constructed a graphical-mathematical theory of competition (MacArthur 1972). MacArthur based his theory on the relationship between population growth rates and resource availability. For example, if species 1 and 2 have abundances X_1 and X_2 depending on resources R_1 and R_2 , then you can use mathematics to describe how the rates of change of X_1 and X_2 depend on the amount of resources R_1 and R_2 . With this information for multiple species, one can predict the resource abundances at which two species can stably coexist.

David Tilman extended MacArthur's theory and developed the R^* , or *Resource-*

Ratio, Theory to provide a mechanistic understanding of competition⁸⁸ (Tilman 1980, 1982). R^* theory is used to predict and explain species-resource correlations. Assume species' growth rates and the resource availability are interdependent. As a population grows, resource availability will limit population size. If species 1 and 2 utilize the same resource, then whichever species can continue to grow on less of the resource should out compete the other. For every species-resource pair, there is a minimum amount of the resource, R^* , which will allow the species to maintain its population size. Coexistence between two species is possible if there are two limiting resources and neither species has a higher R^* for both resources.

Tilman began this work for his dissertation (Tilman 1977, Tilman 1976, Titman 1976)⁸⁹. He tested his theory's predictions with experiments using algae species from Lake Michigan. He measured the R^* 's for a species-resource pair in chemostats, a laboratory environment in which resource levels can be controlled over time. When he competed these species against each other with both resources, he confirmed R^* theory's predictions.⁹⁰ Tilman continues to develop and test R^* Theory, especially at Cedar Creek, a University of Minnesota field station, one of the 26 stations in the Long Term Ecological Research (LTER) Network.

Tilman was an undergraduate at the University of Michigan when the Zoology Department hired Stephen Hubbell in 1969. Tilman has said that he had been interested

⁸⁸ Another way competition theory has gone follows Chesson's coexistence theory. See (Chesson 1994, 2000). R^* theory is a special case of coexistence theory, as is the neutral theory (Adler, HilleRisLambers, and Levine 2007). See (Munoz and Huneman 2016) for a discussion of how the neutral theory is less mechanistic than coexistence theory.

⁸⁹ Tilman changed his name from Titman in 1976 (Sprugel 2015).

⁹⁰ See (Miller et al. 2005) for a recent review.

in studying physics, but while taking his first biology course from Hubbell, Hubbell impressed upon him that ecology was a promising science for theorists (Davis 2004, Sprugel 2015). Tilman stayed at Michigan to do his PhD in zoology. He did part of the work on competition experiments in chemostats with Hubbell. Hubbell had just finished his PhD on energy flow through ecological systems, where he argued for a shift in perspective from one “outside” the organism to one “inside” (Hubbell 1969, Hubbell 1971). He built a series of more complex cybernetic models based on his new foundation.⁹¹ Hubbell moved from Michigan to the University of Iowa in 1976 and published several theoretical papers extending competition theory (Hsu, Hubbell, and Waltman 1977, Hsu, Hubbell, and Waltman 1978, Hansen and Hubbell 1980). But Hubbell soon transitioned from chemostats and competition theory to tropical rain forests and neutral theory.

Hubbell debuted his census data from a 13.44 ha tropical dry forest in Costa Rica together with the ancestor of the neutral theory, the *community drift model*, in 1979 (Hubbell 1979). In 1980, Hubbell co-founded the 50 ha Barro Colorado Island (BCI) Forest Dynamics Plot (FDP) project. The FDP’s primary task is to completely census the plot every 5 years. The first census was completed in 1982, the second in 1985, and every five years since. The BCI FDP has approximately 300 species and between than 210,000 and 250,000 trees and lianas with a diameter at breast height (dbh) greater than or equal to 1 cm. The BCI is now the flagship FDP is the Smithsonian Tropical Research Institute (STRI) Center for Tropical Forest Science - Forest Global Earth Observatory (CTFS-

⁹¹ Interestingly, especially for philosophers of biology, he that Hubbell wrote about (Levins 1966) and model building, especially about generality and realism.

ForestGEO), which now includes over 60 FDPs in forests all over the world. The BCI provides the paradigm case of a neutral community for the neutrality research programme.

The general neutralist approach to resolving the biodiversity paradox is to deny the competitive exclusion principle. Its strategy emphasizes the similarity of species and the role of non-competitive processes such as random birth and death, immigration, and speciation. Where competitionists emphasize the way that the particularities of species determine the species richness and relative abundance of species, neutralists emphasize the similarity of species and the roles of history and chance. If species 1 was more abundant than species 2 in the past, then species 1 will probably continue to be more abundant because it has more members able to fill in any openings left by deaths. Over time, species richness decreases by species going extinct and increases through speciation and immigration.

The view that chance and history shape ecological patterns has a varied history, with three major threads. One thread showcases the MacArthur and Wilson's theory of island biogeography, which modeled species richness on islands as the balance between immigration and extinction (MacArthur and Wilson 1963, 1967). Their emphasis on dispersal processes and neutrality at the species level was taken up, for example, by the Woods Hole Group. Their MBL model of clade diversity was based on speciation and extinction (Raup et al. 1973). Levinton also proposed an ecological model of species

diversity that used speciation and immigration to balance extinction.⁹² Another thread is population biology, where Kimura's neutral theory of molecular evolution synthesized much earlier work and focused later work on neutral dynamics (Kimura 1968, 1983). Caswell imported the neutral theory, gave it an ecological interpretation in terms of speciation, immigration, and drift, and applied it to SADs (Caswell 1976).

A third thread is the equilibrium vs non-equilibrium debate in tropical community ecology. The general problem was whether a tropical rain forest is an equilibrium or nonequilibrium community. An equilibrium community will (i) maintain its SAD when undisturbed and (ii) return to its previous SAD when perturbed. Connell proposed the *equal chance hypothesis* as a particular nonequilibrium verbal model and Hubbell formalized his community drift model in response to it (Connell 1978). Hubbell used the community drift model to undermine the dominant equilibrium explanation of communities, the competitive exclusion principle (Hubbell 1979, Hubbell and Foster 1986). He did this by describing tropical forests SADs using the community drift model. According to the community drift model, patterns resulted of two processes: random drift (birth and death) and random immigration. These processes were random in the sense of being *random with respect to niche differences*, since Hubbell assumed that all individuals in the community had the same chances of giving birth and dying and immigrating. Therefore, if SADs could be fit well without niche differences, these were

⁹² This paper came out the same year, in the same journal as Hubbell's first paper (Hubbell 1979). Hubbell cites it in (Hubbell 2001, Hubbell and Foster 1986) but it seems clear that it was not a direct influence on Hubbell.

Hubbell makes much of the connection to the theory of island biogeography in (Hubbell 1997, 2001) but I argue in Chs 1 and 2 that this is overstated and retrospective.

unnecessary causal factors. Much later, after Hubbell abandoned⁹³ and then took up again working on the community drift model, he incorporated speciation to the model and called it the “unified neutral theory” (Bell 2000, 2001, Hubbell 1997, 2001).

3. Why Scientists are not Falsifying Hypotheses but Developing Research Programmes

Karl Popper’s philosophy of science has been used prominently in ecology since the 1970s (Connor and Simberloff 1979, Roughgarden 1983b, Diamond 1986). But Popper’s philosophy is outdated. In this section I explain a narrative leading from Popper’s falsificationism, through Kuhn’s paradigm based normal science, to Lakatos’ methodology of research programmes. My goal is to develop the weaknesses of the Popperian picture and the strengths of the Kuhnian and Lakatosian philosophies of science. In particular, Popper’s view of science is incapable of making sense of why scientists hold positions such as *Adaptationism* (the view that natural selection is the most important factor in evolution), *Competitionism* (the view that competition is the most important factor in ecology), and the opposing positions to both - namely *Neutralism*. Kuhn’s and Lakatos’ philosophies of science, while opposed on many fronts, share in their opposition to Popper’s strict philosophy of conjecture and refutation. In section 4, I present three versions each of adaptationism, competitionism, and neutralism, and then explain how the scientists developing the neutral theory are committed to

⁹³ Hubbell has stated that the null model wars between Diamond and Simberloff was a significant reason he abandoned work on the community drift model. See chapter 2.

versions of neutralism. Because it can capture neutralism while retaining a Popperian spirit, my view is that something like Lakatos' scientific *research programme* is the best unit of analysis for understanding the continued development of the neutral theory.⁹⁴ This changes how the neutral theory is seen to be used. Rather than falsifying theories based on false predictions, neutralists make predictions in order to make causal conclusions and improve the theory.

3.1 Popper

Karl Popper influenced how scientists and scholars in many different disciplines think about scientific methodology and how science changes over time. It is often said that Popper is the scientists' most popular philosopher of science. This is true especially in ecology, where Popper's ascension coincided both with ecology's fight to become a modern science (see the opening page of (MacArthur 1972) and (Kingsland 1995)). In Popper's image, good science proceeds by the *method of conjecture and refutation*. Scientists propose bold hypotheses and try to refute or falsify their predictions. If they succeed in falsification, they proceed to new conjectures. And if they fail in falsification, they proceed to new attempts at falsification. But, as we will see, conjecture and refutation is far too impoverished to describe how scientists operate over time.

Popper's philosophy of science was an answer to two problems: the Demarcation Problem, and the Problem of Induction (Popper 1963, Ch.1). The Demarcation Problem asks, By what criterion can we demarcate a scientific theory from a pseudo-scientific or

⁹⁴ Philosophers have moved well beyond Kuhn and Lakatos in their understanding of scientific development. But the sequent developments are predicated upon their insights and the story to Kuhn and Lakatos is enough to make my point.

non-scientific theory? Two sorts of theories made a strong impression upon Popper. Theories like Freud's Psychoanalytic Theory were able to explain every possible observation in their domain, while theories like Einstein's General Relativity were only able to explain a limited set of possible observations. Popper's solution to the demarcation problem draws the line between these two sorts of theories. But in order to fully appreciate how and why he answers the demarcation problem as he does, the second problem needs to be understood. Later, it will also show the biggest problem for Popper's falsificationism.

The Problem of Induction asks, How can we justify inductive inferences? One common way of expressing the problem is, Why should we expect the future to resemble the past? To understand the problem and Popper's solution to it, we first need to distinguish several forms of inference.

Deductive inferences are any inferences in which the truth of the premises *guarantees* the truth of the conclusion. *Inductive* inferences do not preserve truth, but rather take the truth of the premises to increase the probability of the truth of the conclusion. The conclusion of an inductive inference is always a stronger claim than the conjunction of the premises. We make two sorts of inductive inferences in scientific reasoning: *inductive generalization* and *abduction*.⁹⁵

Inductive generalization is often used in the context of hypothesis *formation*. An example of an inductive generalization is:

All the swans so far observed have been white.

Therefore, All swans are white.

⁹⁵ This has many other names, including: inference to the best explanation and explanatory inference.

The problem of induction asks for the justification of inductive inferences of this form.⁹⁶ The converse inference is deductive and justified because it preserves truth:

All swans are white.

Therefore, All the observed swans will be white.

For any pair of sentences of the same logical form, it is logically impossible for the premise to be true and the conclusion to be false. But Inductive Generalizations do not preserve truth because there is no contradiction in imagining that the next swan you observe will not be white, no matter how many you have observed. Accepting this form of inductive reasoning can be understood as accepting that the truth of the premises increases, or does not decrease, the probability that the conclusion is true.

Abduction is often used in the context of hypothesis *testing*. The hypothesis that all swans are white can be tested because it deductively implies that the next swan observed will be white. If the next swan observed is not white, then we can deductively reason as follows:

If all swans are white, then the next swan observed will be white.

The next swan observed was not white.

Therefore, not all swans are white.

One counterexample is enough to show that the generalization is false. If the next swan is white, then we can abductively reason as follows:

If all swans are white, then the next swan observed will be white.

The next swan observed was white.

Therefore, all swans are white.

Abduction does not preserve truth because there is no contradiction in imagining that you

⁹⁶ Some deny that induction is formal, but this is a later development. See (Norton 2003).

have merely failed to observe any of the non-white swans that exist. To accept this form of abductive inference is to accept that claims can be confirmed (their probability increased) via testing their deductive consequences. Abduction captures the idea that the best explanation for why the latest observed swan was white is that they all are white, even though, for all you know, they are not.

The Problem of Induction asks for the criterion for good inductive inference and characterizing the forms of inductive inference justified on the criterion. One natural solution is that a good inductive inference is one in which, in any argument of the same form, the truth of the premises increases⁹⁷ the probability of the truth of the conclusion.

Popper solved the Problem of Induction by flatly denying that any inductive inferences are ever justified. His criterion for being a bad inductive inference is just being an inductive inference. He solved the demarcation problem by insisting that scientific theories are able to be refuted or *falsifiable*, while non-scientific theories are unfalsifiable.

Popper argued that any justification for inductive inferences would have to be themselves based on inductive reasoning, since (i) they are either inductive or deductive and (ii) the conclusions of deductive logic are certain but scientific conclusions are never certain. But using induction to justify induction is circular. Therefore, no inductive inferences are justified. For example, if I argue that the past will resemble the future because in the past the future resembled the past, I have merely *assumed* that the future

⁹⁷ Assuming the normal form of inferences. If negations are included, the probabilities need not increase.

will resemble the past and begged the question. For Popper, we have no more reason to infer that the future will resemble the past any more than that it will not resemble the past, no matter how long it has been found to do so.

This is very abstract, but it has practical, prescriptive consequences for how scientists should behave. Because on Popper's view inductive generalization is not justified, scientists are unconstrained by past observations and should conjecture risky, bold hypotheses. And because abduction is not justified, scientists should use deductive logic to attempt to refute or falsify their hypotheses by discovering a false prediction.

Popper proposed that the correct demarcation criterion is that *scientific* theories are *falsifiable*. While many people at the time took theories like Psychoanalysis to have great explanatory power because of their ability to explain every possible observation, Popper instead found them lacking all explanatory power. Theories such as General Relativity, on the other hand, made many prohibitions and specific predictions and for this reason gained explanatory power. For Popper, a theory is scientific just in case it makes prohibitions about what can be observed. This is what it is to be falsifiable. The scientific status of a theory is due to its own character and does not change depending on whether the theory is actually tested and confirmed or refuted. All false theories are falsifiable, and a theory could be falsifiable without ever being actually falsified. This is so far the fate of Einstein's General Theory of Relativity. One of my students remarked that a proper Popperian headline for the recent confirmation of gravitational waves would read, 'Einstein-still not proven wrong!'

3.2 Problems for Popper

While there are undeniably important insights in Popper's solutions, philosophers have in general not been content with conjecture and refutation. For the most part, the demarcation problem has not remained a pressing open problem to be solved. While Popper's falsification criterion has seen important application in issues such as whether intelligent design should be taught in schools (Ruse 1988) and before US courts more generally (Haack 2010), for many philosophers it is no longer seen as an important open problem. Two types of reasons given for abandoning it are that since all attempts have no far proven unsatisfactory, they will probably continue to do so, and that it is a pseudo-problem because there just is nothing shared by all and only scientific theories (Laudan 1983). But many other kinds of proposals have been made and today some still see it as an open problem.

Popper's solution to the problem of induction, on the other hand, has been shown to simply not work. Scientific reasoning requires induction. The basic difficulty, which Popper acknowledged, is that it is not possible to falsify scientific hypotheses without taking for granted other scientific hypotheses. And the obvious way to justify taking scientific hypotheses for granted is citing their record of surviving attempted falsifications. Hence, the denial of induction leads to complete ignorance of both what we know and what we do not know. Assuming we are not completely ignorant, something like confirmation is required for hypothesis testing. This leaves open the interpretation and justification of confirmation.

This death nail for Popper is important enough to the subsequent development of the history of philosophy of science to warrant a thorough explanation. Most scientific theories are not directly falsifiable. They require auxiliary assumptions to derive testable

predictions. Consider Newton's theory of universal gravitation.⁹⁸ The theory consists of Newton's three laws of motion and the inverse-square law. From these hypotheses alone, no observable predictions are derivable. Where do they predict that Mars will be observed? To make actual, testable predictions, auxiliary assumptions are required. For example, reasonable idealized assumptions include (A) there are no bodies except the sun and Mars; (B) there is no friction in the two-body system; (C) Gravitation is the only force acting on the system. Now orbits can be described, but not particular predictions about where to find Mars. Naked eye astronomy still requires optics to work out what we should observe given certain orbits and assumptions about observation errors. And more exact measurements require additional instrumental assumptions including how telescopes work, which depend upon optics to tell us how light behaves when it refracts through our atmosphere.

The basic unit which can be tested then is not the theory or hypothesis. The following is the smallest unit from which testable predictions can be derived:⁹⁹

Central theory

Auxiliary assumptions

Testable prediction

The problem for Popper is that auxiliary assumptions are often themselves general statements which cannot be observed and so require induction to support them. This is a problem because it prevents localizing the false premise(s) in the deductive argument that produced a false prediction. With a deductively valid argument, if the conclusion is false,

⁹⁸ This is the customary example. In what follows I use the following as guides to the auxiliary assumptions and reasoning in the case: (Putnam 1991) (Worrall 2003) (Stanford 2013).

⁹⁹ This was already argued for by Duhem in an older debate (Duhem 1906/1954).

then at least one of the premises must be false, because validity is defined as not taking true premises to false conclusions. If you deduce where Mars should be observed, and something else is observed, should you take this as a refutation of the theory of gravitation? It might be due to a false auxiliary assumption or to a mistake from the instruments. More investigation and reasoning is required to localize the false premise(s).

Popper's solution to this problem was to introduce the concept of *corroboration*: a theory is corroborated to the degree that it survives attempts at falsification. The more corroborated a theory is, the more it can be trusted to not be the source of error in a deduction. The question for Popper then becomes, how is corroboration different from confirmation? Those who accept induction hold that abductive inference increased the confirmation a hypothesis. Popper introduced the concept of corroboration because others interpreted confirmation as having to do with probability of truth. Because Popper denied abduction, his view was that the degree of corroboration had nothing to do with the probability of being *true*, which he always maintained could not be increased by testing. But many interpretations of confirmation are possible. Some of the logical positivists took the goal of science to be trustworthy prediction, not truth, and developed confirmation theory for this goal. Most now see corroboration as another interpretation of confirmation and so an illicit appeal to induction under another name.

Popper's denial of induction is incompatible with the reasoning used in science that Popper himself prescribes. The interpretation of confirmation and probability is remains a live topic of debate. But some theories are should be trustworthy because of their past successful predictions. However, it is still possible to remain Popperian about falsification in the sense that you prescribe the strategy of using tests to falsify scientific

theories and not auxiliary and instrumental assumptions. This is not a trivial position. Thomas Kuhn argued it was normally the wrong strategy for the scientist to adopt.

3.3 Popper vs Kuhn

For Popper, the chief virtue of the scientist is the *critical* attitude of not growing attached to any hypothesis that has so far failed to be refuted. The chief vice of the scientist is being *dogmatic* by growing attached to a favored hypothesis. The official Popperian response to a disagreement between the predictions of their theory and the data is to accept falsification and move on. Doing otherwise makes the theory unfalsifiable and unscientific.

Some genuinely testable theories, when found to be false, are still upheld by their admirers – for example by introducing ad hoc some auxiliary assumption, or by re-interpreting the theory ad hoc in such a way that it escapes refutation. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering its scientific status
(Popper 1963, 48)

Thomas Kuhn proposed a different image of scientific practice centered around the concept of the *paradigm* (Kuhn 1962, 1970b). For Kuhn, scientists are committed to further articulating their paradigm and expanding its reach. Kuhn was famously equivocal in how he used ‘paradigm’, but two main senses stabilized: the paradigm as disciplinary matrix, and the paradigm as exemplar. The paradigm (disciplinary matrix) within which scientists work has least four parts:

- (1) ‘Symbolic generalizations’, including theories and laws such as $f = ma$ or ‘action equal reaction’, which are accepted and to which math and logic can be applied (Kuhn 1962, 182);
- (2) ‘Metaphysical presumptions’, ranging from ontological models such as heat as molecules in

motion, to heuristic models such the billiard-ball model of gases (Kuhn 1962, 184);

(3) 'Values' or what the scientists want from the paradigm. These include the standard virtues of theories such as prediction and explanation (and also what counts as good prediction and explanation) to the more debated virtues such as simplicity and consistency (Kuhn 1962, 185);

(4) The paradigm or 'exemplar' such the two-body problem or harmonic oscillator, which scientists learn in their training and which they rely upon as a shared example of how to solve a problems (Kuhn 1962, 187).

The paradigm-based scientist is a puzzle-solver. The paradigm has a certain empirical domain within which it is hoped it will be able to predict and explain phenomena. The task of the scientist is to expand the actual domain covered by the paradigm by further articulating the theory and expanding the known facts to which it applies. The process of testing hypotheses by deducing predictions from theories and auxiliary assumptions is carried out by scientists, but failed predictions are almost never interpreted as the refutation, or even disconfirmation, of the theories and laws of the paradigm. Instead, these are seen as failures of the scientist to articulate the paradigm in a way that does not fail these tests. As Kuhn said,

[i]n no usual sense, however, are such tests directed to current theory. On the contrary, when engaged with a normal research problem, the scientist must premise current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and his current theory is required to define that puzzle Of course the practitioner of such an enterprise must often test the conjectural puzzle solution that his ingenuity suggests. But only his personal conjecture is tested. (Kuhn 1970a, 4)

Paradigm-based science is called *normal science*. Occasionally puzzles surface which resist solution for a long time acquire the status of an *anomaly*. These find increased attempts as resolution by especially young scientists. A famous example of an

anomaly for the Newtonian paradigm was the discrepancy between the Uranus' predicted orbit and its observed orbit. Many astronomers worked on this problem and many solutions were explored. Some Newtonian entertained slightly modifying the inverse-square law of gravitation. Others used the law of gravitation to calculate where a hypothetical eighth planet should be found if it was causing the discrepancy. Astronomers Adams and Leverrier used this strategy to produce predictions which led to the observation of Neptune. This solution localized the source of the anomaly to the auxiliary assumption that there were only seven planets in the system, saving and expanding the actual scope of the Newtonian paradigm.

Sometimes anomalies are not resolved and the paradigm enters a *crisis*. In a crisis, the norms of normal science are suspended, by some in some situations, and *extraordinary science* begins. Here scientists begin to question whether the puzzle can be solved using the resources of the paradigm. There are two ways of moving beyond a crisis: either the crisis is resolved because the old paradigm proves capable, or the crisis leads to a *scientific revolution* in which a new paradigm is adopted.

Another famous anomaly for the Newtonian paradigm was the precession of the perihelion (the rate at which the ellipse which describes a planet's orbit itself rotates) of Mercury. Astronomers had successfully predicted the processions of the perihelia of all the known planets using Newton's laws and the law of gravitation *except* for Mercury. Eventually this became a crisis. Some theorists again postulated a ninth planet, Vulcan, and calculated where it should be. Nothing was found. Others modified Newton's laws of gravitation from the inverse square law to something slightly different. Yet others

modified Newton's laws themselves.¹⁰⁰ Einstein did this when he produced General Relativity, which finally predicted Mercury's precession successfully. This was a scientific revolution in physics and eventually led to the subordination of the Newtonian paradigm to the Einsteinian.

Popper's scientists use auxiliary assumptions in order to test and hopefully falsify a theory. Kuhn's normal scientists try to solve the puzzle of which auxiliary assumptions, or the observational and experimental material, will enable a theory to make successful predictions. They also depend on the continued existence of new and open puzzles and anomalies (Worrall 2003, 77). From Popper's perspective, Kuhn's normal scientists are dogmatic and in denial. But from Kuhn's perspective, Popper's heroic scientists call for revolution whenever a problem emerges.

Kuhn's normal scientists work to develop the paradigm. This helps us see why an account of how well the evidence supports one theory vs another is not enough to inform scientists about what paradigm they should be working in. Two theories of a given domain can be at drastically different developmental stages. The fact that scientists working on a theory currently do not know any auxiliary assumptions from which to derive successful predictions might say very little about whether this puzzle will be solved in even the near future. Instead of testing hypotheses, scientists are appraisers of the hope and promise of a paradigm.

3.4 Lakatos

¹⁰⁰ For other reasons, this is what Einstein did when he proposed Special Relativity. SR can be seen as an extension of the Newtonian paradigm. GR however cannot.

Lakatos sought a way to incorporate Kuhn's insights into how normal science functions into a normative, Popperian image of how scientists confront theories with data that also solves the demarcation problem. The main unit of analysis for Lakatos is the *research programme*. A research programme very similar to a paradigm. It contains the goals sought, the methods used, the metaphysical commitments assumed, and theories and hypotheses tested. There are two relevant differences between Lakatos' and Kuhn's views of science.

First, Lakatos saw that there are often multiple competing research programmes. Kuhn distinguished mature from immature science on the basis of the acceptance of a common paradigm. This enabled puzzle-solving normal science. But Kuhn further argued that at any given time, a given domain would have one paradigm within which all scientists worked. The question for the scientist within a paradigm is whether to continue normal science or doubt the promise of the paradigm and begin to develop a new one. Lakatos' possibility of multiple research programmes allows for much richer relationships between domains.

This leads to the second relevant difference. Lakatos, like Popper and unlike Kuhn, was not content to let scientists sort out whether a given research programme was worth pursuing. Lakatos accepted that scientists do and should reject their auxiliary assumptions before their core theories. But he went further by distinguishing good from bad ways in which a hypothesis can be saved from falsification by tinkering with the auxiliary assumptions. To explain this, he introduced some terminology that I will follow.

Call the central theories and hypotheses of a research programme its *hard core*. Every research programme follows a *negative heuristic* which directs scientists to not

modify the hard core in response to a failed test of a theory's prediction. Call the auxiliary and instrumental assumptions of a research programme its *protective belt*. Every research programme also follows a *positive heuristic* to modify the protective belt when predictions fail. So far this is no different than Kuhn said normal scientists do in paradigms. There are broadly two ways in which the negative heuristic can be carried out: the hard core can be saved from falsification in a way that makes *novel predictions*, or in an *ad hoc* way which does not. Then, a research programme is *progressive* if it tends to respond to negative tests in ways that predict novel facts which are then confirmed. A research programme is *degenerate* if it avoids false predictions in continually ad hoc ways.

Lakatos' solution to the demarcation problem then goes as follows. First, the question becomes whether the *method* or *activity* is scientific. His view is that progressive research programmes are scientific, and degenerate ones are not. Good science is about stating ahead of time what conditions would conflict with a theory, and progressive research programmes do this by making clear predictions about novel facts. Lakatos is a falsificationist in the sense that he states that a theory T is falsified by another theory T' when T' predicts more novel facts than T, grows around T, and some excess of T' is corroborated (Lakatos 1976, 32).

Lakatos is also intended his solution to the demarcation problem to guide scientists' behavior. Therefore, he should prescribe scientists to abandon a degenerate research programme and follow a progressive one. The main problem for Lakatos is that this is too strong, but that nothing weaker will do. Lakatos admits that research programmes often go through degenerate phases, especially in their beginning, before

developing progressively. The past success of the research programme alone is not a guide to its promise or future success. This depends upon the world, the research programme, and the scientists working to develop it further. It is ironic that the problem for a denier of induction in scientific reasoning is that he cannot induct over history either. This means Lakatos can issue no strong norms for when to abandon a research programme or when a research programme is irredeemably degenerate.

Going forward, I draw the following lessons. First, confirmation is required for scientific reasoning. This means that inductive inference and support for theories and hypotheses via confirmation is warranted and should be not critiqued on the grounds that the study did not falsify any hypotheses. Second, a failed prediction is not a reason in itself to abandon a theory. Retaining a theory after a failed test is not necessarily behaving unscientifically. Third, the way in which scientists retain theories and modify them in response to failed predictions is best captured by focusing on something like the paradigm or research programme. This unit contains core theories, but it also includes the goals, methods, systems studied, and paradigm examples of well-solved problems which guide the development of auxiliary assumptions. Fourth, research programmes can exist side by side in the same domain and compete with each other.

4. Adaptationism, Competitionism, Neutralism

I began calling the proponents and developers of the neutral theory *neutralists*, in contrast to the competitionists, in section 2. In this section I characterize *Neutralism*, and use this to characterize the *neutralist research programme* in section 5. Neutralism is an

alternative to *Competitionism*, which is the ecological analog of *Adaptationism*.

Adaptationism is the position that natural selection is the most important factor in evolution. Anti-adaptationists challenged this position across evolution, paleobiology, and ecology starting in the 1970s (Nitecki and Hoffman 1987). In all three disciplines, neutral theories were developed, called ‘null’, and used to counter adaptationism.

Neutralism also exists in evolution as one alternative to adaptationism. Evolutionary neutralism developed around Kimura’s neutral theory of molecular evolution. Both evolutionary adaptationism and neutralism have been characterized in terms of research programmes, and this should be extended competitionism and neutralism.

4.1 Adaptationism

Three kinds of adaptationism can be distinguished: empirical adaptationism, explanatory adaptationism, and methodological adaptationism (Godfrey-Smith 2001, Orzack and Forber Winter 2012). *Empirical adaptationism* is the view that natural selection is the strongest cause of evolution. Most traits can be predicted and explained by taking into account only natural selection. This is a position on what Beatty (Beatty 1997) has characterized as a *relative significance problem*. A relative significance problem asks which of a number of processes, or causal factors, is most strongly or more frequently, or both, the cause of a pattern. Empirical adaptationism is the contingent and empirical claim that natural selection dominates other causes of evolution. An alternative position along this dimension of adaptationism might emphasize the role of developmental constraints or neutral drift.

Explanatory adaptationism is the view that (i) adaptations are the most important

aspects of the biological world to be explained, and that (ii) natural selection is the best explanation of adaptations. This concerns both what it is most important to explain and what type of explanation is most important. Adaptations here include both intrinsic organization of organisms and the fit between an organism and their environment. Explanatory adaptationists want to explain apparent design in organisms, and do so using adaptation via natural selection. This is independent of empirical adaptationism because it may be held by someone who accepts that natural selection is not the dominant process of evolution in general (Godfrey-Smith 2001, 336). This is a non-empirical claim about how to organize research, the goals of evolutionary science, and/or aesthetics (Orzack and Forber Winter 2012). An alternative position along this dimension might focus on growth and development as the most important feature to be explained, and emphasize the role of phylogeny to explain development.

Methodological adaptationism is the view that looking for adaptations and signatures of natural selection are the most effective strategies for approaching evolutionary research.¹⁰¹ Ernst Mayr advocated this view,

“[The evolutionary biologist] must first attempt to explain biological phenomena and processes as the product of natural selection. Only after all attempts to do so have failed, is he justified in designating the unexplained residue tentatively as a product of chance.” (Mayr 1983) Quoted in (Beatty 1987)

Methodological positions such as this direct research. An alternative position along this dimension might instead direct research to begin with chance. This has been expressed strongly by Robert Selander,

¹⁰¹ Sarkar argues that explanatory and methodological adaptationism are hard to distinguish in practice (Sarkar 2014). I think this is often true, but in the case of ecology it is useful to distinguish them.

“All of our work begins with tests of the null hypothesis that variation in allele frequencies generated by random genetic drift is the primary cause of molecular evolutionary change. This is the logical point of departure. ... In our view, natural selection should be invoked only when the stochastic model is rejected.” (Selander 1985, 87) Quoted in (Beatty 1987)

The language of ‘testing the null hypothesis’ is always used by methodological positions contrary to adaptationism or its analogs in other areas of biology.

This tripartite distinction highlights three of the different kinds of commitments of a paradigm or research programme. Research programmes all contain empirical, explanatory, and methodological positions. A given adaptationist or rival research programme may mix and match different kinds and strengths of adaptationist positions.¹⁰² Methodological adaptationism in particular highlights the main feature of paradigms and research programmes which Popper’s falsificationism gets wrong - the tenacity of holding to a core theoretical claim.

Consider how Gould¹⁰³ critiqued the adaptationist programme. Gould objected to the way that adaptationists continue to seek explanations for traits in terms of natural selection even when none have so far been found. The strongest form of this reasoning is the prescription: “If one adaptive argument fails, try another.” (Gould and Lewontin 1979, 586). Gould criticized adaptationists for not accepting the failure of an adaptive

¹⁰² Some directly call adaptationism a research programme. I think this is wrong because there is no unique adaptationist research programme. For example, multiple programmes might all push methodological adaptationism, but with different theories and models and goals. The following discuss adaptationism in the context of research programmes:

Scientists: (Gould and Lewontin 1979, Mayr 1983, Mitchell and Valone 1990)

Philosophers: (Sober 1993, Ch5, Orzack and Sober 1994, Resnik 1996, Sterelny and Griffiths 1999, Godfrey-Smith 2009, Pievani 2012, Allegra 2011-2012)

¹⁰³ I say Gould because Lewontin has said in an interview that Gould wrote most of it, except what I interpret to mean section 5. See <https://evolution-institute.org/article/the-spandrels-of-san-marco-revisited-an-interview>

hypothesis as reason to look for non-adaptive explanations. But this is just what one expects when thinking in terms of research programmes. Gould's is not a valid criticism because it depends upon the outdated Popperian framework. It also mis-characterizes how an adaptationist programme operates. Particular adaptationist hypotheses may be rejected as wrong, but others will be subsequently proposed and evaluated.

4.2 Competitionism

Community ecology studies the patterns of species that are found in the same place and that interact with each other. The domain of community ecology lies between population ecology, which studies single population growth, and ecosystem ecology, which studies the full set of interactions between the biotic populations and abiotic conditions in a place. While community ecology can embrace interactions such as predation, much of it limits itself to the dynamics of single trophic level communities, such as the plants in a tropical rain forest. Patterns that community ecologists aim to explain include resource-trait/species patterns, species coexistence patterns, biodiversity patterns, and biogeographical patterns.

To explain these patterns, ecologists draw on the relation between evolution and ecology. Over evolutionary time, species are adapted to environments by natural selection and other processes. Over ecological time, species come to occupy a particular environment by interspecific competition and other ecological processes. In this way, interspecific competition is the ecological analog of natural selection and is made by possible by it. Put another way, natural selection and other evolutionary processes determine the *fundamental niche*, the set of conditions within which life is possible, of a

species (Hutchinson 1957). Interspecific competition and other ecological processes determine the *realized niche*, the conditions within which a species is actually living, of a species. In the same way that a species can be better *adapted* evolutionarily to one environment than another, a species can have an ecological *good fit* to an environment, as when, for example, the species are stable in the environment. The connection between adaptedness and fit, and selection and competition, gives rise to the ecological analogs of adaptationism. I denominate this family of positions *Competitionism*.¹⁰⁴ Three kinds of competitionism can be distinguished.

Empirical competitionism is the view that interspecific competition and competitive niche exclusion are the most important causal processes shaping the realized niches of populations in a community. It is a position on what testing hypotheses has already revealed about the relative significance question, and on what future tests will probably reveal.

Explanatory competitionism is the view that (i) good fits between species and environment is the most important feature of the ecological world to be explained, and that (ii) interspecific competition is the best explanation of good fits. The idea that good fits are to be explained has two parts. One feature is the fit between resources and species. The other feature is the fit between species and other species. Interspecific competition is the classic explanation of how two species whose fundamental niches overlap and whose actual distributions overlap come to have their stable realized niches.

¹⁰⁴ As far as I know, no one has named the analog of adaptationism or its associated research programme. But they have come close.

Strong wrote, "Competition in ecology has attributes of a paradigm (sensu Kuhn 1962)" (Strong 1980, 280). Schoener wrote about the 'competitionist' views and paradigm (Schoener 1982).

The competitive exclusion principle describes the strongest way that this can happen, when one species completely excludes the other because they use a single limiting resource. When there are multiple limiting resources, coexistence is possible when neither species has a higher R^* for both resources.

The competitive exclusion principle is described in terms of *stable coexistence*, and the idea of stability is important to competitionism. One feature expected in a community dominated by competition is *stability*¹⁰⁵ - the ability to maintain the current assemblage of species and their relative species abundances without disturbance and to move back to the stable state after a disturbance. So a community in which all the species fit the environment and each other well is expected to be stable over time.

Interspecific competition is also used to explain species coexistence, biodiversity, and biogeography patterns, although the resource abundances are not taken into account here. Jared Diamond's explanation of the coexistence patterns of bird species on islands with interspecific competition is a paradigm of this kind of explanation (Diamond 1975). Here some species have competitive advantages over others and so come to exclude one another on different islands.

Methodological competitionism is the view that looking for good fit between organisms and their environment and signatures of interspecific competition are the most effective strategies for approaching ecological research. The competitive exclusion

¹⁰⁵ There is a huge and varied ecological literature on stability. Many meanings have been distinguished and studied theoretically and empirically.

The stability-diversity hypothesis is that more species rich communities are more stable. This fits right in the middle of the biodiversity paradox. It is unclear if the stability-diversity hypothesis is true. See (May 1973) for a classic work, and (Ives and Carpenter 2007) for more recent. See (Tilman, Reich, and Knops 2006) for work done by Tilman on grasslands.

principle is used by methodological competitions to investigate previously undetected niche differences.¹⁰⁶ In order to resolve the biodiversity paradox, this is carried out until relevant species differences are found, or the species are shown to not stably coexist. In this way the principle operates as an organizing principle and gives direction to research.

Just as in evolution, a counter strategy was put forward in community ecology. Instead of assuming interspecific competition and related mechanisms, Simberloff and Strong, leaders of the Tallahassee Mafia at FSU, wrote,

“... we propose another possibility with logical primacy over other hypotheses, that other hypotheses must first be tested against, but that is rarely considered at all by ecologists. This is the null hypothesis that community characteristics are apparently random ... “ (Strong Jr, Szyska, and Simberloff 1979, 910)

The Simberloff and Strong continued to deploy and develop their alternative strategy.

The important point is that it is opposed to methodological competitionism and proposes an alternative starting point for analyzing ecological communities.¹⁰⁷

4.3 Ecological Neutralism

Three kinds of *neutralisms*, opposed to competitionism and based on non-competitive processes, can now be distinguished. One issue for characterizing neutralism

¹⁰⁶ For a great exchange centered on how the competitive exclusion principle is used to give explanations, see: (Weber 1999), (Eliot 2011), and (Raerinne and Baedke 2015)

¹⁰⁷ The Tallahassee Mafia and the Woods Hole Group waver between treating their ‘null hypotheses’ of ‘randomness’ as purely statistical hypotheses and substantial hypotheses based on chancy processes. This leads them to waver between what I distinguish in chapter 4 as *null modeling* and *baseline modeling*.

is that neutrality is conflated with demographic and dispersal processes.¹⁰⁸ But they are distinct, though not independent. Neutrality is the assumption that all individuals in a community, regardless of species, have identical ecological probabilities. Assuming neutrality excludes interspecific competition, because interspecific competition requires differences between species. One way to think of the neutral model is that it portrays what interspecific competition would be like if it was the same as *intraspecific* competition - competition for resources (space) among equals. All models with limited resources include this kind of competition.

The processes of drift, immigration, and speciation are lumped together as demographic dispersal processes. The neutral model assumes neutrality and these three processes, but any combination of demographic and dispersal processes might be included in a neutral model. Also, any combination of demographic and dispersal processes might be included in a non-neutral model, which would then also include interspecific competition. Because competition is based upon non-neutrality, there are uncountably-many competitive hypotheses.

The neutral model also assumes constant, or approximately constant,¹⁰⁹ finite community size. If an infinite-size community model were considered, it would experience no drift, because drift comes about through competition for space (or random sampling). The simplest case here excludes immigration and speciation. If the infinite-

¹⁰⁸ Ecologist Jeremy Fox has called this a ‘zombie idea’ on his blog: <https://dynamicecology.wordpress.com/2012/01/23/zombie-ideas-in-ecology-neutral-stochastic/>

¹⁰⁹ Hubbell originally assumed constant community size, but then it was shown that the same theorems follow from assuming that the community size randomly walks about the same size.

size model is neutral, then the community will stochastically maintain its initial SAD unless disturbed. If the infinite-size model is non-neutral, then it will evolve from its initial state to a stable state. The infinite-size models can be further augmented to include immigration and speciation, in which new expectations can be derived.

A neutral-demographic-dispersal community is a quasi-stable community. A neutral community will not in general retain the same species over time over regain them after a disturbance. In this sense they are unstable. But neutral-demographic-dispersal communities do approach and approximately maintain a stable SAD over time. If many species were removed from a community, over time the species richness is expected to increase from speciation and immigration. In this sense they are stable.

Neutralism is opposed to competitionism because it assumes neutrality, not demographic or dispersal processes. But *demographic-dispersalism* could also be defined. As we will see, the neutralist is committed to both neutrality and demographic and dispersal processes in similar, defeasible, ways. In what follows I characterize neutralism in terms of *neutrality*, but keep in mind that neutrality without demographic or dispersal processes would make a very boring model.

Empirical neutralism is the view that the competitive differences between individuals of different species are causally less important than the similarities between individuals in a community. This means that the community is more neutral than competitive with respect to the patterns of interest. Empirical neutralism is also associated with asserting that the processes of random drift, immigration/ dispersal, and speciation are more important than competitive processes shaping patterns of abundance

and diversity. Together, it asserts that neutral demographic and dispersal processes causally dominate the patterns of interest. For the most part, empirical neutralism is about the same patterns as competitionism. However, neutralism tends to focus on patterns of abundance and diversity. In particular, trait-resource patterns and species co-occurrence are an interesting issue for the neutralist. These patterns are generally still taken to be within the explanatory scope of empirical neutralism, but because neutralism ignores species differences, the neutralist predictions often fail.

Explanatory neutralism is the view that (i) the patterns of abundance and diversity are the patterns to be explained, independent of whether there is good fit of species to environment, and (ii) neutrality and demographic and dispersal processes are the best explanation of these processes. This is the hardest neutralist position to carve out. With respect to the first part, I am not aware of anyone having expressed such a position in ecology (although I do not doubt someone has) so I am extrapolating from what Peter Godfrey-Smith has said about how Kimura argued against explanatory adaptationism and argued that not only the patterns of adaptive genes but also neutral genes. In this sense, community ecology has for a long time included in its explanatory target patterns of abundance and diversity which do not concern species-environment fit directly. They have just usually proposed competitive explanations of these patterns. With respect to the second part, again I am not aware of an ecologist arguing that neutral demographic and dispersal explanations are inherently better than others at explaining some set of patterns. At this point, neutral-demographic-dispersal and competitive processes are taken to be

good explanations of these patterns.¹¹⁰

Methodological neutralism is the view that determining the strength of various processes targeted for explanation is best approached by starting with the position that they are caused by neutral demographic and dispersal processes. Future research can build upon this starting point by relaxing the neutrality assumption. It can also modify other assumptions such as frequency-independence, which says that the growth rate is independent of the population size. Frequency-dependence is typically associated with competitive explanations because it is a stabilizing process. As mentioned in the cases of methodological adaptationism and competitionism, explanatory neutralism has also been expressed using “null hypothesis” terminology. Hubbell writes,

What community-level patterns result from the simplifying assumption that all species in a community exhibit the same demographic stochasticity (ecological drift) on a per capita basis, and ecological drift is the only process occurring besides random dispersal and speciation? ...

The value of posing this question rigorously is that we obtain a quantitative null hypothesis against which to test when, to what extent, and for which species demographic differences among species are necessary to explain observed community patterns. (Hubbell 2006, 1387)

The problem with this, as I have argued extensively in chapters 3 and 4, is that “null hypothesis” mischaracterizes the reasoning being followed.

Going forward, the distinctions between the different kinds of competitionism will be used to characterize the neutralist research programme and primary dimension of disagreement between the competitionist and neutralist research programmes. But first,

¹¹⁰ This is true despite the rhetoric. See chapter 4 for an extensive discussion of how ecologists argue in terms of *theory choice questions* but actually address *relative*

the neutralist research programme needs to be characterized.

5. The neutralist research programme

5.1 The problem context

I begin my explaining the problem context which the neutralist and competitionist research programmes address. The basic problem is, “What are the patterns of community abundance and diversity and what processes produce those patterns?” There are many strategies for answering this question. I am focusing on one modeling strategy used by the neutralists. There are five basic steps:

(1) *Gather abundance data.* A single-trophic level community is either completely censused or sampled and the species of every individual is identified and recorded. The data consists of a list of species and the number of individuals in each species. A single-trophic level community consists of all of the organisms at the same energy level, such as all the plants in a forest or all of the nesting birds on an island except for the raptors. This excludes interactions between trophic levels such as predation and disease.

significance questions. Beatty has argued that this is characteristic of many biological arguments (Beatty 1997).

(2) *Represent the abundance data.* There are many patterns of abundance.¹¹¹ The most common way of representing abundance data is to plot it on a graph. This permits quick visual inspection and comparison across communities. There are multiple ways of representing the data and they do not all contain the same information. One representation method is called a Whittaker or dominance-diversity plot. Figure 5.1 is a comparison of five communities' dominance-diversity plots:

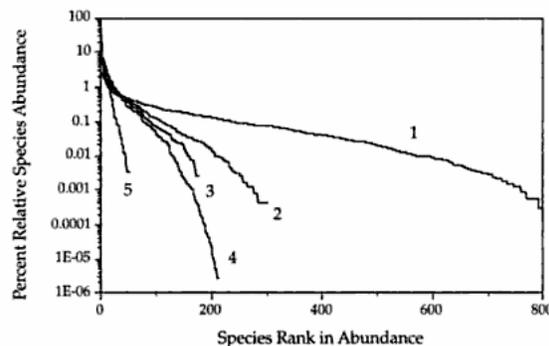


FIG. 1.1. Patterns of relative species abundance in a diverse array of ecological communities. Species in each community are ranked in percentage relative abundance from commonest (*left*) to rarest (*right*). The percentage relative abundance is log transformed on the y-axis. 1: Tropical wet forest in Amazonia. 2: Tropical dry deciduous forest in Costa Rica. 3: Marine planktonic copepod community from the North Pacific gyre. 4: Terrestrial breeding birds of Britain. 5: Tropical bat community from Panama.

Figure 5.1 – Comparison of species abundance distributions
From (Hubbell 2001, 4). Used with permission from Princeton University Press.

The y-axis represents the percent relative abundance of a species and the x-axis represents the percent relative abundance of a species, both of which are often plotted on a \log_{10} scale. Each line represents a community and a point on a line represents a species

¹¹¹ I just focus on SADs in this chapter, because they are the main pattern discussed in the context of neutral theory and the reasoning patterns are not different for different patterns.

The kinds of patterns about which neutral theory makes predictions is greater than SADs and one area of development of the theory is expanding its predictions to include, for example, spatial and temporal abundance patterns. The holy grail would be dynamical spatial and temporal abundance distribution modeling, but this is not there yet.

and is read as ‘the x th ranked species makes up $y\%$ of the community’. Dominance-diversity plots make the evenness of a community very apparent; the tropical rainforest has many more species of roughly the same abundance than the tropical bat community.

A dominance-diversity plot, and not the original abundance data, is a species abundance distribution (SAD). A dominance-diversity plot contains less information than the original abundance data because it does not include the names of the different species. This facilitates comparison of SADs across communities which contain different species. Other forms of SADs can contain less information but make other aspects of communities more apparent.¹¹²

An example of a regularity in the data that community ecologists want to explain are the facts that in almost all SADs, most species are relatively rare and few are common (McGill et al. 2007, 996).

(3) *Construct statistical models of representations of abundance data.* Predicting and explaining the patterns of abundance require describing the patterns by fitting them to a statistical model. A statistical model is just something which outputs a statistical distribution whose shape is tuned via values of its free parameters. Fisher’s logseries distribution is an early example of a statistical model constructed for and applied to SADs (Fisher, Corbet, and Williams 1943). But causally explaining patterns using models requires the model to have a causal interpretation. MacArthur was the first to do this in ecology, when he constructed the *broken-stick distribution* to represent the competitive

¹¹² See (Magurran 2004, McGill et al. 2007) for other forms of SADs and their virtues and vices.

exclusion principle (MacArthur 1960, MacArthur 1957).¹¹³ In this case, claims of the theory are formalized into mathematics, the mathematics and auxiliary assumptions are used to construct a statistical model.

(4) *Compare statistical models with representations of abundance data.* The fit between a statistical model and a SAD can be measured using statistical techniques and is called the *model-fit* (Lloyd 1994, Lloyd 2010). Model-fit relative to a SAD can either be determined for a given set of values of free parameters, or the best possible model-fit can be determined by finding the set of values which maximize the model fit.¹¹⁴ Robustness or sensitivity analysis is also done to determine the dependency of parameter values on the model-fit.

(5) *Respond to model-fit.* Depending on the goals and standards set, model-fits vary from excellent to poor. Judgments about whether the model-fit is good enough for a particular purpose are then made and acted on. The typical responses of the research programme are described as heuristics after the neutral theory and neutral models are characterized.¹¹⁵

¹¹³ It is worth noting that MacArthur abandoned this project fairly quickly (MacArthur 1966).

¹¹⁴ Though their paper is about introducing the Akaike information criteria, it is very helpful for explaining how statistical model-fit is measured. See (Forster and Sober 1994).

¹¹⁵ This chapter avoids getting into the justification for the heuristics, but perhaps the strongest criticism of the neutrality research programme is that good model-fits for SADs are easy to come by and give no relative confirmation to neutral theory over other theories which produce the same patterns. This is more than an underdetermination argument. It is primarily about data like SADs just not having very much information.

5.2 The neutral theory and neutral model

The neutral theory is a theory of the large-scale structure in the patterns of the kinds of species in a community. If you look at a forest on the scale of a tropical forest on an island in the Panama Canal to the whole Amazon Basin, the patterns that stand out include the number of species and the number of individuals in each species. For example, McGill has called the fact that many SADs (including those in the figure above) have an S-shape one of ecology's few descriptive laws (McGill et al. 2007, 996). Because SADs have a similar shape across many communities, they probably share a common cause. The common causes of communities at this scale are three processes: *ecological drift*, interpreted as random birth and death; *immigration*, the source of diversity for the local community such as the birds on an island; and *speciation*, the source of diversity on the larger immigrant pool such as the mainland. These dispersal processes change the community via influencing the probability that a given open space will be filled by an individual of a given species. Given this, over time a given species' relative abundance will appear to randomly walk up and down in abundance. The only limiting resource in a community is space and the size of a community stays close to constant.

In what follows, the neutral model presented in (Hubbell 2001) is the "basic neutral model". It is the basis for most of the subsequent work done developing the theory. The following is the ecological interpretation of the model, which is formalized using mathematics. Imagine a community as a checkerboard, each space represents some place for an individual to live. Fill the board with pieces of various colors drawn from a bag. The board represents the island or local community and the bag represents the

mainland or metacommunity. The neutral theory models how the SAD changes over time and predict what the equilibrium SAD is expected to be based on several interpreted free parameters. Each turn consists of removing a piece and replacing the space with a new piece to maintain constant community size. To begin, remove a piece at random (i.e. with equal probability for each piece), this represents a death. The death can either be replaced by a local birth or by immigration. With probability $1-m$, replace the death with a piece according to the following rule: each piece currently on the board has an equal chance of reproducing and filling the space. Therefore, the probability that a given color will fill the space is proportional to the current relative abundance of that color. This follows from assuming *neutrality*, that all individuals in the community are functionally equivalent. With probability m , replace a death according to the following rule: each piece currently in the bag has an equal chance of immigrating and filling the space. For this the dynamics of a metacommunity must also be tracked, and these follow analogous rules. Here drift occurs and is balanced by speciation, instead of immigration, where a death in the bag is replaced by an individual of a novel species. Eventually, an equilibrium (time-independent) SAD will be reached.

The basic model has three free, interpreted parameters: J , local community size; m , immigration rate; and θ , a function of J_M , metacommunity size, and v , speciation rate. The model predicts that the equilibrium patterns will be statistically indistinguishable from a particular statistical distribution. Differences in SADs in different communities and time are caused by differences in these parameter values.

5.3 The heuristics of the neutral theory of ecology

The neutral dispersal research programme responds to model-fit measurements in four ways. If there is no deviation, then it judges that the neutral theory “properly describes what matters most while ignoring irrelevant details” - Heuristic 1. If there is deviation, then either it is a problem with the theory or model, or else there is no problem. If it is a problem, then either it has to do with the way the theoretical assumptions were formalized - Heuristic 2, or it has to do with the theoretical assumptions themselves - Heuristic 3. Finally, if it is not a problem, then the neutral model is detecting powerful forces beyond the baseline neutral-demographic-dispersal - Heuristic 4. All together, these heuristics allow constant work to be done in the research programme on the important problems. I illustrate each heuristic with a brief example of its use.

Heuristic 1: Abduction. One goal of the neutralist research programme is to give causal explanations of community patterns. Neutralists draw causal conclusions using abductive inference when there is good model-fit between a neutral model and a SAD. An example of such a conclusion is the following:

“Species interactions, niche partitioning, or density-dependence, while they may be present, do not appear to enhance tree species richness at Barro Colorado.” (Condit, Chisholm, and Hubbell 2012a, 1)

This is a *negative conclusion* about the lack of responsibility of several processes traditionally associated with competition theory in the BCI FDP. *Positive conclusions* are also drawn,

“The neutral model predicts diversity and abundance at Barro Colorado because it properly describes what matters most – species input [neutral speciation] – while ignoring irrelevant details.” (Condit, Chisholm, and Hubbell 2012a, 5)

Hubbell claims that the *burden of proof* is on those who push some competitionist explanation of abundance patterns to first show that the neutral theory is inadequate.

“...obtaining acceptable fits from neutral models shifts the burden of proof to those who would assert that more complex theory is required to explain nature and with what level of detail and generality” (Hubbell 2006, 1387)

And until the burden of proof is met, Hubbell gives neutralist explanations the benefit of the doubt in the sense that they are accepted until rejected.

Heuristic 2: Adding Complexity to the Model. If the neutral theory is accepted, statistical measurement is unchallenged, and there is deviation, then the problem must be in the mathematical model. Ricklefs (Ricklefs 2003, 2006, Leigh 1981) tested the neutral theory against the average life-spans of species in tropical forest plots. Hubbell had proposed two distinct models of speciation, as a point mutation or as a random fission, in (Hubbell 2001) and done the math for the point mutation.¹¹⁶ Ricklefs showed the point mutation model consistently overestimates extinction rates (and underestimates species life spans) while the random fission model consistently underestimates extinction rates (and overestimates species life spans).

In response, Hubbell (Hubbell 2003) modified an auxiliary assumption about how speciation worked. He invoked an intermediate form of speciation, *peripheral isolate* speciation, where new species are produced with a small but greater than 1 population size, which he had first proposed in (Hubbell and Lake 2003). This intermediate form can vary the size of novel populations, and is capable of fitting extinction rate data better than

¹¹⁶ Much of this math was already worked out in population genetics for Kimura’s neutral theory.

either other model.¹¹⁷

Heuristic 3: Adding Complexity to the Theory. A core theoretical claim of the neutralist is that population growth is density-*independent*, because neutrality makes birth rate independent of species and independent of current abundance. Many studies have argued that some form of density-dependence is important for regulating abundance and diversity in communities. This has been true, in particular, for studies of tropical forest communities.¹¹⁸ Volkov (Volkov et al. 2005) incorporated density-dependence into the neutral theory via the equivalent but more tractable mathematical framework developed in (Volkov et al. 2003).

This augmented neutral theory relaxes the assumption that growth rate is independent of current relative abundance but keeps the assumption that growth rate is independent of species. Hubbell (Hubbell 2008) claims that satisfying the latter alone is sufficient for being *neutral*. But the view of community dynamics changes. Rare species are uncommon on the density-independent view because most of them go extinct very quickly due to drift. But rare species are uncommon on the density-dependent view because being rare gives rare species a boost in growth rate that makes them more common. So the latter view requires more than drift and immigration and is not a strictly dispersal view, even if it is neutral.

¹¹⁷ This back and forth has continued and several other modes of speciation have been formalized and incorporated into the neutral model. The most extensive survey of this development is (Kopp 2010). Kopp points out that the peripheral isolate model of speciation has still not been tested.

¹¹⁸ See, for example: (Leigh Jr 1999) (Chesson 2000) (Henderson and Magurran 2014). But also see (Condit, Hubbell, and Foster 1994).

Both forms of neutral models are capable of good model-fits and are statistically indistinguishable from each other on the SADs of six tropical forest plots. Because the density-independent neutral model is simpler, hypothesizing density-dependence in these communities is said to be unnecessary. Comparing these two models is further used to estimate the strength of density-dependence using the fourth heuristic. The analysis shows that any rare species advantage only occurs in very small populations.

Many other examples for this case can be given. Environmental stochasticity, another way that diversity is lowered, was included into the theory (Benedetti-Cecchi 2007, Alonso, Etienne, and McKane 2007, Kalyuzhny, Kadmon, and Shnerb 2014, Kalyuzhny et al. 2014). The neutrality assumption was relaxed to form the nearly-neutral theory (Zhou and Zhang 2008) (He, Zhang, and Lin 2012).

These cases share the common feature of a claim that was previously given the status of a core claim has been relaxed. This suggests that rather than a strict dichotomy between core claims and auxiliary assumptions, there are more and less core claims which are protected from being modified according to how deep entrenched they are.

Heuristic 4: Baseline Modeling. With the benefit of the doubt is given to the neutral theory, another heuristic explains deviations as signs of additional causes not included in the current neutral theory. For example, Terborgh et al. (Terborgh, Foster, and Nuñez 1996) evaluated the neutral theory on five mature floodplain forest plots along the Manu river in Peru. They argued that the correlation between the abundances of the most common species across disconnected forest plots was too unlikely to be produced by the neutral model (Hubbell 2001, 330, Terborgh, Foster, and Nuñez 1996, 564).

Hubbell responded that, if it is assumed that the five sites are all connected by immigration and so are not independent, then the neutral theory predicts that the most abundant species in one local community will be the commonest species in another local community “essentially 100% of the time” (Hubbell 2001, 335). This neutralized the problem.

Additionally, Hubbell argued that the correlation in which the commonest species dominate the other species in abundance was *excessive*. This case of *ecological dominance deviation* revealed, Hubbell estimated, a 6% fitness advantage in the commonest species (Hubbell 2001, 337), a clear deviation from neutrality.

This the same pattern of reasoning that Adams and Leverrier used to find Uranus and save the Newtonian research programme from falsification. Another episode of the baseline modeling heuristic used an augmented neutral model to argue that the increase in diversity following the breakup of Pangaea was not due to continental drift and isolation alone (Jordan, Barraclough, and Rosindell 2016). They did this by showing that the predictions of the neutral deviated significantly from the observations.

5.4 The explanatory-developmental dialectic

These four heuristics work in concert to guide the explanatory and developmental practice of the neutrality research programme over time. The following list is a way of see how this happens. At stake is the debate between explanatory competitionism and explanatory neutralism. This relative significance question is answered by studying various communities and determining the relative significance of the multiple processes. For this reason, steps 1 and 2 may refer to different data from different communities. I

use “neutral model*” because the process can be iterated with any version of a neutral model.

1. Neutralists abduct from good model-fit between neutral model* and pattern P (SADs, species lifespans, etc.) in communities C to conclusion that some part of the neutral theory* explains what causes pattern P in communities C at a time and place and scale (See Heuristic 1.)
2. Challengers argue that the model-fit between the neutral model* and pattern P in other communities D is poor.¹¹⁹ They take this as *disconfirming evidence* for the neutral theory and in the strongest form claim to have refuted the neutral theory (Dornelas, Connolly, and Hughes 2006).
3. Proponents respond in one of two ways:
 4. *Baseline Modeling*: Deviations between the neutral model* and pattern P are a way to detect and measure the presence and strength of causal processes and factors not included in the neutral theory causally responsible for pattern P. (Heuristic 4.)
 5. *Adding complexity*: Neutralists respond that poor model-fit between neutral model* and some P is reason to augment the neutral model* with additional complexity. This can take the form of either:
 6. Changing the way that a process or factor is formalized in the neutral model*. Example: peripheral isolate speciation. (Heuristic

¹¹⁹ Again, for this chapter I am bracketing the objection that model-fits to patterns such as SADs are basically uninformative.

2.)

7. Including additional processes or causal factors in the neutral theory and model. Example: density-dependence, near-neutrality, environmental stochasticity. (Heuristic 3.)
8. This reasoning is iterated with the same data or new data until either step 1 or 3.1 are arrived at.

The strategies of *Baseline Modeling* and *Adding Complexity* function to avoid disconfirming evidence leading to falsification. Baseline modeling turns poor model-fit into evidence of the dominance of competitive processes. Adding complexity turns poor model-fit into evidence that they have not yet made the neutral theory and neutral models good enough.

6. The debate between competitionists and neutralists

What are the axes of disagreement between competitionists and neutralists? Both groups aim to solve a *relative significance problem* over the relative contributions of competition and neutral-dispersal processes to patterns in community ecology. But both groups accept that their basic theory and models are incomplete and expect that sometimes the processes which they say are dominant are not. This debate between empirical competitionism and empirical neutralism is ongoing, and will only be decided in the future.

The primary disagreement is between methodological competitionism and methodological neutralism over the best research strategy for investigating the relative

significance problem. Settling the debate between empirical competitionism and empirical neutralism depends upon each research programme developing their respective theories to the best of their abilities. The neutralist starts with neutral-dispersal processes and sometimes relaxes these, either by relaxing neutrality or by including non-dispersal processes or both. Hubbell has expressed this very clearly:¹²⁰

“I argue, seemingly paradoxically, that the most rapid path to understanding ecological systems, especially complex systems such as species-rich tropical tree communities, is through an interaction of empirical science, guided by strong inference, with theories that start very simply, with few free parameters and assumptions, and add complexity reluctantly, kicking and screaming, only when absolutely necessary to obtain some desired level of fit to the data. Neutral theory is one such starting point.” (Hubbell 2008, 143)

The competitionist research programme takes the opposite approach. It starts with strong interspecific interactions for resources and relaxes this to include dispersal processes or lower the strength of competition (for example, (Tilman 2004)). Both use their false, incomplete models (Wimsatt 1987) to advance their research programmes from their starting point using heuristics.

Many small-scale (chemostats to small plots such as Cedar Creek’s) systems seem to be dominated by interspecific competition. These systems can be controlled to allow testing of hypotheses about how resource levels correlate with species richness and abundance. Many intermediate-scale systems (such as tropical forests and coral reefs) seem to be dominated by neutral-dispersal processes. These systems cannot be controlled

¹²⁰ The use of “strong inference” is telling of the Popperian influence on ecology still. This is a reference to (Platt 1964) who argued for Popperian falsificationism and testing multiple hypotheses against each other via crucial experiments.

and experiments are impossible. The biggest difference from the debate between adaptationists and neutralists in evolution is that in ecology there will never be the diversity and quality of data for community abundance and diversity as there is for DNA sequences because of the differences in scale and available control. In the face of this complexity, the competitionists and neutralists have taken different paths.

There are tradeoffs between the two paths. Tilman (Stanley Harpole and Tilman 2006) thinks it is a weakness of the neutral theory that it will not ever satisfactorily explain species-resource correlations. Hubbell (Hubbell 2010) thinks it is a weakness that any adequate R^* theory will require several free parameter per species-resource pair in a community (remember the BCI FDP has 300 species of freestanding trees). But the potential explanatory domain of both approaches partially coincide and so a decision must be made about from which to generalize. This decision needs to be informed by the best-developed versions of each theory available.

There is much more at stake and much more going in community ecology on than testing hypotheses. But because of the background views on methodology from philosophy of science and statistics, issues such as strategies for dealing with poor model-fits and whether the goal is understanding or control are not discussed as openly as testing hypotheses is. But goals and methods are just as central to science as hypotheses and worldviews.

A background assumption in the debate between competitionists and neutralists is that *there can be only one*. This analog of competitive exclusion at the level of research programmes is *meta-methodological monism*. Monists hope and expect a day when there is one research programme to rule them all. Perhaps the neutral theory, R^* theory, a new

theory, or some hybrid theory that includes all kinds of processes will be at the center of it.¹²¹ But as interspecific competitive exclusion, inter-research programme competitive exclusion can also be resisted. *Meta-methodological pluralism* allows that one or the other may or may not come to exclude the other in the future. It is not at all obvious that when there are multiple goals, methods, and worldviews in play, any direct effort should be made cull all but one research programme standing.

7. Conclusion

In this chapter I have shown how the neutral theory's explanatory practices, development, and relationship to competition theory are illuminated by focusing on the unit of the research programme. I have done this by first arguing that the history of philosophy of science shows that a Popperian view of science as conjecture of refutation of hypotheses and theories is too impoverished to account for both inductive inference in science and the development of research programmes. I next argued that the three kinds of adaptationism philosophers of biology developed for evolutionary biology can be extended to characterize three forms each of competitionism and neutralism. I then characterized the heuristics used by the neutralist research programme to explain using and develop the neutral theory. Finally, I characterized the competitionist research programmes as primarily representing different methodological starting points and paths

¹²¹ Vellend (Vellend 2010) has pushed for the hybrid view and includes drift, immigration, speciation, and competition in the framework. But he ignores or does not comment on whether this new synthetic theory will adopt methodological competitionism or neutralism, or some other position. This is the same problem as with Lloyd's (Lloyd 2015) *multiple factors view* of evolution, which does not give any *strategy* for deciding what to do with disconfirming evidence.

to understanding ecological systems. This analysis will not settle many of the disagreements between the neutralists and competitionists. But it is not meant to. My hope is that characterizing the neutral theory and research programme will help to focus attention on the further articulations and justifications for the neutralist and competitionist research programmes towards the goals of greater understanding the discipline of ecology and the ecological world.

References

- Adler, Peter B, Janneke HilleRisLambers, and Jonathan M Levine. 2007. "A niche for neutrality." *Ecology letters* 10 (2):95-104.
- Allegra, Alessandro. 2011-2012. "Assessing Methodological Adaptationism: an Historical Approach." MSc, Philosophy, London School of Economics.
- Alonso, David, Rampal S Etienne, and Alan J McKane. 2007. "Response to Benedetti-Cecchi: Neutrality and environmental fluctuations." *Update* 22 (5).
- Beatty, John. 1984. "Chance and natural selection." *Philosophy of Science*:183-211.
- Beatty, John. 1987. "Natural selection and the null hypothesis." In *The latest on the best*, edited by John Dupre. Cambridge: MIT Press
- Beatty, John. 1997. "Why do biologists argue like they do?" *Philosophy of Science*:S432-S443.
- Bell, Graham. 2000. "The distribution of abundance in neutral communities." *The American Naturalist* 155 (5):606-617.
- Bell, Graham. 2001. "Neutral macroecology." *Science* 293 (5539):2413-2418.
- Benedetti-Cecchi, Lisandro. 2007. "Neutral theory and 1/f noise make similar predictions of assemblage dynamics." *Trends in ecology & evolution* 22 (5):231.
- Brown, James H, and Brian A Maurer. 1989. "Macroecology: the division of food and space among species on continents." *Science* 243 (4895):1145-1150.
- Cadotte, Marc. 2015. "The first null model war in ecology didn't prevent the second one." *The EEB & Flow*, July 17, 2015. <http://evol-eco.blogspot.ca/2015/07/the-first-null-model-war-in-ecology.html>.
- Call, Josep, and Michael Tomasello. 2008. "Does the chimpanzee have a theory of mind? 30 years later." *Trends in cognitive sciences* 12 (5):187-192.
- Caswell, Hal. 1976. "Community structure: a neutral model analysis." *Ecological monographs* 46 (3):327-354.
- Chase, Jonathan M, and Mathew A Leibold. 2003. *Ecological niches: linking classical and contemporary approaches*: University of Chicago Press.
- Chen, Youhua. 2016. "Local speciation can be incorporated into neutral theory of biodiversity." *Ecological Modelling* 325:67-70.
- Chesson, Peter. 1994. "Multispecies competition in variable environments." *Theoretical Population Biology* 45 (3):227-276.
- Chesson, Peter. 2000. "Mechanisms of maintenance of species diversity." *Annual review of Ecology and Systematics*:343-366.
- Christie, Thony. 2016. "Isaac and the apple – the story and the myth." *The Renaissance Mathematicus*. <https://thonyc.wordpress.com/2016/05/12/isaac-and-the-apple-the-story-and-the-myth/>.
- Cohen, Joel E. 1968. "Alternate derivations of a species-abundance relation." *American naturalist*:165-172.
- Condit, Richard, Ryan A Chisholm, and Stephen P Hubbell. 2012a. "Thirty years of forest census at Barro Colorado and the importance of immigration in maintaining diversity." *PloS one* 7 (11):e49826.
- Condit, Richard, Ryan A Chisholm, and Stephen P Hubbell. 2012b. "Thirty years of forest census at Barro Colorado and the importance of immigration in maintaining diversity." *PloS one* 7 (11).
- Condit, Richard, Stephen P Hubbell, and Robin B Foster. 1994. "Density dependence in two understory tree species in a neotropical forest." *Ecology* 75 (3):671-680.
- Connell, Joseph H. 1978. "Diversity in tropical rain forests and coral reefs." *Science* 199 (4335):1302-1310.

- Connor, Edward F, and Daniel Simberloff. 1979. "The assembly of species communities: chance or competition?" *Ecology*:1132-1140.
- Cooper, Gregory. 1993. "The competition controversy in community ecology." *Biology and Philosophy* 8 (4):359-384.
- Darwin, Charles. 1859. "The origin of species." *London: Murray*.
- Davis, Tinsley H. 2004. "Biography of David Tilman." *Proceedings of the National Academy of Sciences of the United States of America* 101 (30):10851-10853. doi: 10.1073/pnas.0404605101.
- de Mazancourt, Claire. 2001. "Consequences of Community Drift." *Science* 293 (5536):1772. doi: 10.1126/science.1064207.
- Diamond, Jared. 1986. "Overview: laboratory experiments, field experiments, and natural experiments." In *Community ecology*, edited by Jared Diamond and Ted J Case, 3-22.
- Diamond, Jared M. 1975. "Assembly of species communities." In *Ecology and evolution of communities*, edited by Martin Cody and Jared M Diamond, 342-444. Cambridge, Massachusetts: Harvard University Press.
- Dienes, Zoltan. 2008. *Understanding psychology as a science: An introduction to scientific and statistical inference*. New York: Palgrave Macmillan.
- Dietrich, Michael R. 2008. "Molecular Evolution." In *A Companion to the Philosophy of Biology*, 157-168. Blackwell Publishing Ltd.
- Dornelas, Maria, Sean R Connolly, and Terence P Hughes. 2006. "Coral reef diversity refutes the neutral theory of biodiversity." *Nature* 440 (7080):80-82.
- Douglas, Heather. 2009. *Science, policy, and the value-free ideal*: University of Pittsburgh Pre.
- Duhem, Pierre. 1906/1954. *The aim and structure of physical theory*. Translated by Philip P. Wiener. Princeton: Princeton University Press.
- Eigen, Manfred, and Ruthild Winkler. 1981. *Laws of the game: how the principles of nature govern chance*. New York: Knopf.
- Eliot, Christopher H. 2011. "Competition theory and channeling explanation." *Philosophy & Theory in Biology* 3.
- Elton, Charles. 1946. "Competition and the structure of ecological communities." *The Journal of Animal Ecology*:54-68.
- Elton, Charles S. 1927. *Animal ecology*: University of Chicago Press.
- Etienne, Rampal S, and David Alonso. 2007. "Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence." *Journal of Statistical Physics* 128 (1-2):485-510.
- Etienne, Rampal S, and Bart Haegeman. 2011. "The neutral theory of biodiversity with random fission speciation." *Theoretical Ecology* 4 (1):87-109.
- Ewens, Warren J. 1972. "The sampling theory of selectively neutral alleles." *Theoretical population biology* 3 (1):87-112.
- Fisher, Ronald A, A Steven Corbet, and Carrington B Williams. 1943. "The relation between the number of species and the number of individuals in a random sample of an animal population." *The Journal of Animal Ecology*:42-58.
- Forster, Malcolm, and Elliott Sober. 1994. "How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions." *The British Journal for the Philosophy of Science* 45 (1):1-35.
- Fox, Jeremy. 2012. "Zombie ideas in ecology: "neutral" = "dispersal limitation". " *Dynamical Ecology*, January 24 2012. <https://dynamicecology.wordpress.com/2012/01/24/zombie-ideas-in-ecology-neutral-dispersal-limitation/>.
- Gause, George Francis. 1934. *The struggle for existence*: Williams & Wilkins Co. .

- Gause, Georgii Frantsevich. 1932. "Experimental studies on the struggle for existence I. Mixed population of two species of yeast." *Journal of experimental biology* 9 (4):389-402.
- Giere, Ronald N. 1998. *Explaining science: A cognitive approach*: University of Chicago Press.
- Gigerenzer, Gerd. 2004. "Mindless statistics." *The Journal of Socio-Economics* 33 (5):587-606.
- Gilpin, Michael E, and Jared M Diamond. 1984. "Are species co-occurrences on islands non-random, and are null hypotheses useful in community ecology." In *Ecological communities: conceptual issues and the evidence.*, edited by Donald R Strong Jr, Daniel Simberloff, Lawrence G Abele and Anne B Thistle, 297-315. Princeton, New Jersey, USA: Princeton University Press,.
- Godfrey-Smith, Peter. 1994. "Of Nulls and Norms." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.
- Godfrey-Smith, Peter. 2001. "Three kinds of adaptationism." *Adaptationism and optimality*:335-357.
- Godfrey-Smith, Peter. 2009. *Theory and reality: An introduction to the philosophy of science*: University of Chicago Press.
- Gotelli, Nicholas J, and Gary R Graves. 1996. *Null models in ecology*. Washington DC: Smithsonian Institution Press.
- Gotelli, Nicholas J, and Brian J McGill. 2006. "Null versus neutral models: what's the difference?" *Ecography* 29 (5):793-800.
- Gould, Stephen Jay, and Richard C Lewontin. 1979. "The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme." *Proceedings of the Royal Society of London B: Biological Sciences* 205 (1161):581-598.
- Grinnell, Joseph. 1917a. "Field tests of theories concerning distributional control." *American Naturalist*:115-128.
- Grinnell, Joseph. 1917b. "The niche-relationships of the California Thrasher." *The Auk*:427-433.
- Haack, Susan. 2010. "Federal Philosophy of Science: A Deconstruction-And a Reconstruction." *NYUJL & Liberty* 5:394.
- Hansen, Stephen R, and Stephen P Hubbell. 1980. "Single-nutrient microbial competition: qualitative agreement between experimental and theoretically forecast outcomes." *Science* 207 (4438):1491-1493.
- Hare, Brian, Josep Call, and Michael Tomasello. 2006. "Chimpanzees deceive a human competitor by hiding." *Cognition* 101 (3):495-514.
- He, Fangliang, Da-Yong Zhang, and Kui Lin. 2012. "Coexistence of nearly neutral species." *Journal of plant ecology* 5 (1):72-81.
- Henderson, Peter A, and Anne E Magurran. 2014. "Direct evidence that density-dependent regulation underpins the temporal stability of abundant species in a diverse animal community." *Proceedings of the Royal Society of London B: Biological Sciences* 281 (1791):20141336.
- Hsu, SB, SP Hubbell, and Paul Waltman. 1978. "A contribution to the theory of competing predators." *Ecological Monographs*:337-349.
- Hsu, Sze-Bi, S Hubbell, and Paul Waltman. 1977. "A mathematical theory for single-nutrient competition in continuous cultures of micro-organisms." *SIAM Journal on Applied Mathematics* 32 (2):366-383.
- Hubbell, Stephen P. 1971. "Of sowbugs and systems: the ecological bioenergetics of a terrestrial isopod." *Systems analysis and simulation in ecology* 1:269-324.
- Hubbell, Stephen P. 1979. "Tree dispersion, abundance, and diversity in a tropical dry forest." *Science* 203 (4387):1299-1309.

- Hubbell, Stephen P. 1984. "Methodologies for the study of the origin and maintenance of tree diversity in tropical rainforest." In *The Significance of Species Diversity in Tropical Forest Ecosystems*, edited by Gema Maury-Lechon, Malcom Hadley and Talal Younes, 8-17. International Union of Biological Sciences.
- Hubbell, Stephen P. 1995. "Towards a theory of biodiversity and biogeography on continuous landscapes." *Progress in biometeorology*.
- Hubbell, Stephen P. 1997. "A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs." *Coral reefs* 16 (1):S9-S21.
- Hubbell, Stephen P. 2001. *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Vol. 32: Princeton University Press.
- Hubbell, Stephen P. 2003. "Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs." *Oikos* 100 (1):193-199.
- Hubbell, Stephen P. 2004. "Two decades of research on the BCI forest dynamics plot." In *Tropical forest diversity and dynamism: findings from a large-scale plot network*, edited by Elizabeth Claire Losos and Egbert G Leigh, 8-30. Chicago, Illinois, USA: University of Chicago Press.
- Hubbell, Stephen P. 2005. "The neutral theory of biodiversity and biogeography and Stephen Jay Gould." *Paleobiology* 31 (sp5):122-132.
- Hubbell, Stephen P. 2006. "Neutral theory and the evolution of ecological equivalence." *Ecology* 87 (6):1387-1398.
- Hubbell, Stephen P. 2008. "Approaching ecological complexity from the perspective of symmetric neutral theory." In *Tropical forest community ecology*. Wiley-Blackwell, Chichester, UK, edited by Walter Carson and Stefan Schnitzer, 143-159. Wiley.
- Hubbell, Stephen P. 2010. "Neutral theory and the theory of island biogeography." In *The theory of island biogeography revisited*, edited by Jonathan B. Losos and Robert E Ricklefs, 264-292. Princeton, New Jersey: Princeton University Press.
- Hubbell, Stephen P, and Robin B Foster. 1982. *The Structure and Dynamics of a Tropical Forest: A Demographic Analysis*. The University of Iowa: National Science Foundation. Original edition, DEB-8206992.
- Hubbell, Stephen P, and Robin B Foster. 1986. "Biology, chance, and history and the structure of tropical rain forest tree communities." In *Community ecology*, edited by Jared Diamond and Ted J Case, 330. Harper & Row.
- Hubbell, Stephen P, and Jeffrey K Lake. 2003. "The neutral theory of biodiversity and biogeography, and beyond." *Macroecology: concepts and consequences*. Blackwell, Oxford:45-63.
- Hubbell, Stephen Philip. 1969. "A systems analysis of the ecological bioenergetics of a terrestrial isopod." University of California, Berkeley.
- Huss, John Edward. 2004. "Experimental reasoning in non-experimental science: Case studies from paleobiology." University of Chicago, Committee on the Conceptual and Historical Studies of Science.
- Huss, John Edward. 2009. "The shape of evolution: the MBL model and clade shape." In *The paleobiological revolution: Essays on the growth of modern paleontology*, edited by David Sepkoski and Michael Ruse, 568. University of Chicago Press.
- Hutchinson, G Evelyn. 1957. "Concluding remarks." *Cold spring harbor symposium on quantitative biology* 22:415-427.
- Hutchinson, G Evelyn. 1961. "The paradox of the plankton." *American Naturalist*:137-145.
- Ives, Anthony R, and Stephen R Carpenter. 2007. "Stability and diversity of ecosystems." *science* 317 (5834):58-62.

- Jarvinen, Olli. 1982. "Species-to-genus ratios in biogeography: a historical note." *Journal of Biogeography*:363-370.
- Jordan, Sean MR, Timothy G Barraclough, and James Rosindell. 2016. "Quantifying the effects of the break up of Pangaea on global terrestrial diversification with neutral theory." *Phil. Trans. R. Soc. B* 371 (1691):20150221.
- Kalyuzhny, Michael, Ronen Kadmon, and Nadav M Shnerb. 2014. "A generalized neutral theory explains static and dynamic properties of biotic communities." *arXiv preprint arXiv:1406.2926*.
- Kalyuzhny, Michael, Efrat Seri, Rachel Chocron, Curtis H Flather, Ronen Kadmon, and Nadav M Shnerb. 2014. "Niche versus neutrality: a dynamical analysis." *The American Naturalist* 184 (4):439-446.
- Karlin, Samuel, and James McGregor. 1972. "Polymorphisms for genetic and ecological systems with weak coupling." *Theoretical population biology* 3 (2):210-238.
- Kimura, Motoo. 1968. "Evolutionary rate at the molecular level." *Nature* 217 (5129):624-626.
- Kimura, Motoo. 1983. *The neutral theory of molecular evolution*: Cambridge University Press.
- Kimura, Motoo, and Tomoko Ohta. 1977. "Further comments on "counter-examples to a neutralist hypothesis"." *Journal of molecular evolution* 9 (4):367-368.
- Kingsland, Sharon E. 1995. *Modeling nature*: University of Chicago Press.
- Kopp, Michael. 2010. "Speciation and the neutral theory of biodiversity." *Bioessays* 32 (7):564-570.
- Kuhn, Thomas S. 1962. *The structure of scientific revolutions*: University of Chicago press.
- Kuhn, Thomas S. 1970a. "Logic of discovery or psychology of research." In *Criticism and the Growth of Knowledge*, edited by Imre Lakatos and Alan Musgrave, 1-23. New York: Cambridge University Press.
- Kuhn, Thomas S. 1970b. *The structure of scientific revolutions*. 2nd ed: University of Chicago press. Original edition, 1962.
- Lack, David Lambert. 1976. *Island biology: Illustrated by the land birds of Jamaica*. Berkeley and Los Angeles: University of California Press.
- Lakatos, Imre. 1976. *Falsification and the methodology of scientific research programmes*: Springer.
- Laudan, Larry. 1983. "The demise of the demarcation problem." In *Physics, philosophy and psychoanalysis*, 111-127. Springer.
- Leigh, Egbert Giles. 1981. "The average lifetime of a population in a varying environment." *Journal of Theoretical Biology* 90 (2):213-239.
- Leigh, Egbert Giles. 2007. "Neutral theory: a historical perspective." *Journal of evolutionary biology* 20 (6):2075-2091.
- Leigh Jr, Egbert Giles. 1999. *Tropical Forest Ecology: A View from Barro Colorado Island: A View from Barro Colorado Island*: Oxford University Press.
- Levins, Richard. 1966. "The strategy of model building in population biology." *American scientist*:421-431.
- Levinton, Jeffrey S. 1979. "A Theory of Diversity Equilibrium and Morphological Evolution." *Science* 204 (4390):335-336. doi: 10.1126/science.204.4390.335.
- Lloyd, Elisabeth A. 2010. "Confirmation and robustness of climate models." *Philosophy of Science* 77 (5):971-984.
- Lloyd, Elisabeth A. 2015. "Adaptationism and the Logic of Research Questions: How to Think Clearly About Evolutionary Causes." *Biological Theory*:1-20.
- Lloyd, Elisabeth Anne. 1994. *The structure and confirmation of evolutionary theory*: Princeton University Press.

- Longino, Helen E. 2008. "Values, heuristics, and the politics of knowledge." In *The challenge of the social and the pressure of practice, science and values revisited*, 68-86.
- Losos, Elizabeth Claire, and Egbert Giles Leigh. 2004. *Tropical forest diversity and dynamism: Findings from a large-scale plot network*: University of Chicago Press.
- Lotka, Alfred J. 1925. *Elements of physical biology*.
- Lurz, Robert W. 2011. *Mindreading animals: the debate over what animals know about other minds*. Cambridge: MIT Press.
- MacArthur, Robert. 1960. "On the relative abundance of species." *American Naturalist*:25-36.
- MacArthur, Robert. 1966. "Note on Mrs. Pielou's Comments." *Ecology* 47 (6):1074.
- MacArthur, Robert H. 1957. "On the relative abundance of bird species." *Proceedings of the National Academy of Sciences of the United States of America* 43 (3):293.
- MacArthur, Robert H. 1972. *Geographical ecology: patterns in the distribution of species*: Princeton University Press.
- MacArthur, Robert H, and Edward O Wilson. 1963. "An equilibrium theory of insular zoogeography." *Evolution*:373-387.
- MacArthur, Robert H, and Edward O Wilson. 1967. *The theory of island biogeography*. Vol. 1: Princeton University Press.
- Magurran, Anne E. 1988. *Ecological diversity and its measurement*: Springer Netherlands.
- Magurran, Anne E. 2004. *Measuring biological diversity*: Blackwell Science.
- May, Robert M. 1975. "Patterns of species abundance and diversity." In *Ecology and evolution of communities*, edited by Martin Cody and Jared Diamond, 81-120. Harvard University Press.
- May, Robert McCredie. 1973. *Stability and complexity in model ecosystems*. Vol. 6: Princeton University Press.
- Mayo, Deborah G, and Aris Spanos. 2006. "Severe testing as a basic concept in a Neyman–Pearson philosophy of induction." *The British Journal for the Philosophy of Science* 57 (2):323-357.
- Mayr, Ernst. 1983. "How to carry out the adaptationist program?" *American Naturalist*:324-334.
- McGill, Brian. 2003. "Strong and weak tests of macroecological theory." *Oikos* 102 (3):679-685.
- McGill, Brian J, Rampal S Etienne, John S Gray, David Alonso, Marti J Anderson, Habtamu Kassa Benecha, Maria Dornelas, Brian J Enquist, Jessica L Green, Fangliang He, Allen H Hurlbert, Anne E Magurran, Pablo A Marquet, Brian A Maurer, Annette Ostling, Candan U Soykan, Karl L Ugland, and Ethan P White. 2007. "Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework." *Ecology letters* 10 (10):995-1015.
- Melis, Alicia P, Josep Call, and Michael Tomasello. 2006. "Chimpanzees (Pan troglodytes) conceal visual and auditory information from others." *Journal of Comparative Psychology* 120 (2):154.
- Miller, Thomas E, Jean H Burns, Pablo Munguia, Eric L Walters, Jamie M Kneitel, Paul M Richards, Nicolas Mouquet, and Hannah L Buckley. 2005. "A critical review of twenty years' use of the resource-ratio theory." *The American Naturalist* 165 (4):439-448.
- Millstein, Roberta L. 2000. "Chance and macroevolution." *Philosophy of Science*:603-624.

- Millstein, Roberta L, Robert A Skipper Jr, and Michael R Dietrich. 2009. "(Mis) interpreting mathematical models: drift as a physical process." *Philosophy & Theory in Biology* 1.
- Mitchell, William A, and Thomas J Valone. 1990. "The optimization research program: studying adaptations by their function." *Quarterly Review of Biology*:43-52.
- Mori, Terutaka. 2013. "The discovery of species–abundance distribution in an ecological community." *Oikos* 122 (2):179-182.
- Munoz, François, and Phillipe Huneman. 2016. The neutral theory of community ecology: is it null?
- Nekola, Jeffrey C, and James H Brown. 2007. "The wealth of species: ecological communities, complex systems and the legacy of Frank Preston." *Ecology letters* 10 (3):188-196.
- Neyman, Jerzy, and Egon S Pearson. 1933. "The testing of statistical hypotheses in relation to probabilities a priori." *Mathematical Proceedings of the Cambridge Philosophical Society*.
- Nitecki, Matthew H, and Antoni Hoffman. 1987. *Neutral models in biology*: Oxford University Press.
- Norton, John D. 2003. "A Material Theory of Induction*." *Philosophy of Science* 70 (4):647-670.
- Odenbaugh, Jay. Forthcoming. "Neutrality, Niche, and Nulls: Causal Relevance in Ecology." In *Philosophical Perspectives on Causal Reasoning in Biology*, edited by C Kenneth Waters and James Woodward. Minneapolis: University of Minnesota Press.
- Orians, Gordon H. 1962. "Natural selection and ecological theory." *The American Naturalist* 96 (890):257-263.
- Orzack, Steven Hecht, and Patrick Forber. Winter 2012. Adaptationism. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Orzack, Steven Hecht, and Elliott Sober. 1994. "Optimality models and the test of adaptationism." *American Naturalist*:361-380.
- Penn, Derek C, Keith J Holyoak, and Daniel J Povinelli. 2008. "Darwin's mistake: Explaining the discontinuity between human and nonhuman minds." *Behavioral and Brain Sciences* 31 (02):109-130.
- Penn, Derek C, and Daniel J Povinelli. 2007. "On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 362 (1480):731-744.
- Penn, Derek C, and Daniel J Povinelli. 2009. "On becoming approximately rational: the relational reinterpretation hypothesis." In *Rational animals, irrational humans*, edited by Shigeru Watanabe, Aaron P. Blaisdell, Ludwig Huber and Allan Young, 23-43. Tokyo: Keio University Press.
- Penn, Derek C, and Daniel J Povinelli. 2013. "The comparative delusion: the 'behavioristic'/'mentalistic' dichotomy in comparative Theory of Mind research." *Agency and joint attention*:62-82.
- Pievani, Telmo. 2012. "An evolving research programme: The structure of evolutionary theory from a Lakatosian perspective." In *The theory of evolution and its impact*, 211-228. Springer.
- Platt, John R. 1964. "Strong inference." *science* 146 (3642):347-353.
- Popper, Karl. 1963. *Conjectures and refutations: the growth of scientific knowledge*. New York, New York: Harper and Row. Original edition, 1963.
- Povinelli, Daniel J, and Jennifer Vonk. 2004. "We don't need a microscope to explore the chimpanzee's mind." *Mind & Language* 19 (1):1-28.
- Preston, Frank W. 1948. "The commonness, and rarity, of species." *Ecology* 29 (3):254-283.

- Preston, Frank W. 1962. "The canonical distribution of commonness and rarity: Part I." *Ecology* 43 (2):185-215.
- Pueyo, Salvador. 2006. "Diversity: between neutrality and structure." *Oikos* 112 (2):392-405.
- Purves, Drew W, and Stephen W Pacala. 2005. "Ecological drift in niche-structured communities: neutral pattern does not imply neutral process." In *Biotic interactions in the tropics: their role in the maintenance of species diversity*, edited by DFRP Burslem, MA Pinard and SE Hartley, 107-138.
- Putnam, Hilary. 1991. "The "corroboration" of theories." In *The philosophy of science*, edited by Richardson Boyd, Philip Gasper and J.D. Trout, 121-137. Cambridge, Massachusetts: The MIT Press.
- Quinn, James F, and Arthur E Dunham. 1983. "On hypothesis testing in ecology and evolution." *The American Naturalist*:602-617.
- Raerinne, Jani, and Jan Baedke. 2015. "Exclusions, Explanations, and Exceptions: On the Causal and Lawlike Status of the Competitive Exclusion Principle." *Philosophy & Theory in Biology* 7.
- Raup, David M. 1987. "Neutral models in paleobiology." In *Neutral models in biology*, 121-132. New York: Oxford University Press.
- Raup, David M, and Stephen Jay Gould. 1974. "Stochastic simulation and evolution of morphology-towards a nomothetic paleontology." *Systematic Biology* 23 (3):305-322.
- Raup, David M, Stephen Jay Gould, Thomas JM Schopf, and Daniel S Simberloff. 1973. "Stochastic models of phylogeny and the evolution of diversity." *The Journal of Geology*:525-542.
- Resnik, David. 1996. "Adaptationism: hypothesis or heuristic?" *Biology and Philosophy* 12 (1):39-50.
- Ricklefs, Robert E. 2003. "A comment on Hubbell's zero-sum ecological drift model." *Oikos* 100 (1):185-192.
- Ricklefs, Robert E. 2006. "The unified neutral theory of biodiversity: do the numbers add up?" *Ecology* 87 (6):1424-1431.
- Rosindell, James, Stephen P Hubbell, and Rampal S Etienne. 2011. "The unified neutral theory of biodiversity and biogeography at age ten." *Trends in ecology & evolution* 26 (7):340-348.
- Roughgarden, Jonathan. 1983a. "Competition and theory in community ecology." *American Naturalist*:583-601.
- Roughgarden, Jonathan. 1983b. "Competition and theory in community ecology." *The American Naturalist* 122 (5):583-601.
- Ruse, Michael, ed. 1988. *But is it Science? The Philosophical Question in the Creation/Evolution Controversy*. Buffalo, N.Y.: Prometheus.
- Sani, Fabio, and John Todman. 2008. *Experimental design and statistics for psychology: a first course*: John Wiley & Sons.
- Sarkar, Sahotra. 2014. "The genomic challenge to adaptationism." *The British Journal for the Philosophy of Science*:axu002.
- Schoener, Thomas W. 1982. "The controversy over interspecific competition: despite spirited criticism, competition continues to occupy a major domain in ecological thought." *American Scientist*:586-595.
- Selander, Robert K. 1985. "Protein polymorphism and the genetic structure of natural populations of bacteria." In *Population genetics and molecular evolution*, edited by Tomoko Ohta and K. Aoki, 85-106. Tokyo: Japan Scientific Studies.
- Sepkoski, David. 2012. *Rereading the fossil record: the growth of paleobiology as an evolutionary discipline*: University of Chicago Press.
- Simberloff, Daniel S. 1970. "Taxonomic diversity of island biotas." *Evolution*:23-47.

- Simberloff, Daniel S. 1974. "Equilibrium theory of island biogeography and ecology." *Annual review of Ecology and Systematics* 5:161-182.
- Sloep, Peter B. 1986. "Null hypotheses in ecology: towards the dissolution of a controversy." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.
- Sober, Elliott. 1993. "Philosophy of biology."
- Sober, Elliott. 1994. "Let's Razor Ockham's Razor." In *From a biological point of view: Essays in evolutionary philosophy*, 136-157. Cambridge University Press.
- Sober, Elliott. 2002. "Instrumentalism, parsimony, and the Akaike framework." *Philosophy of Science* 69 (S3):S112-S123.
- Sober, Elliott. 2009. "Parsimony and models of animal minds." In *The philosophy of animal minds*, edited by Robert W Lurz, 237. Cambridge University Press.
- Sprugel, Doug. 2015. "G. David Tilman." Ecological Society of America, accessed June 15. <http://esa.org/history/tilman-g-d/>.
- Stanford, Kyle. 2013. Underdetermination of Scientific Theory. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Stanley Harpole, W, and David Tilman. 2006. "Non-neutral patterns of species abundance in grassland communities." *Ecology Letters* 9 (1):15-23.
- Sterelny, Kim, and P Griffiths. 1999. "Sex and death." *Evolution* 22:2.2.
- Strong, Donald R. 1980. "Null hypotheses in ecology." *Synthese* 43 (2):271-285.
- Strong Jr, Donald R, Lee Ann Szyska, and Daniel S Simberloff. 1979. "Test of community-wide character displacement against null hypotheses." *Evolution*:897-913.
- Sugihara, George. 1980. "Minimal community structure: an explanation of species abundance patterns." *American Naturalist*:770-787.
- Suppes, Patrick. 1962. "Models of data." In *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, edited by Ernst Nagel, Patrick Suppes and Alfred Tarski, 252-261. Stanford: Stanford University Press.
- Tansley, AG. 1917. "On competition between *Galium saxatile* L.(*G. hercynicum* Weig.) and *Galium sylvestre* Poll.(*G. asperum* Schreb.) on different types of soil." *The Journal of Ecology*:173-179.
- Terborgh, John, Robin B Foster, and Percy V. Nuñez. 1996. "Tropical tree communities: a test of the nonequilibrium hypothesis." *Ecology* 77 (2):561-567.
- Tilman, David. 1977. "Resource competition between plankton algae: an experimental and theoretical approach." *Ecology* 58 (2):338-348.
- Tilman, David. 1980. "Resources: a graphical-mechanistic approach to competition and predation." *American Naturalist*:362-393.
- Tilman, David. 1982. *Resource Competition and Community Structure. (Mpb-17)*: Princeton University Press Princeton, NJ.
- Tilman, David. 2004. "Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly." *Proceedings of the National Academy of Sciences of the United States of America* 101 (30):10854-10861.
- Tilman, David (published as D. Titman). 1976. "Ecological competition between algae: experimental confirmation of resource-based competition theory." *Science* 192 (4238):463-465.
- Tilman, David, and Peter M Kareiva. 1997. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Vol. 30: Princeton University Press.
- Tilman, David, Peter B Reich, and Johannes MH Knops. 2006. "Biodiversity and ecosystem stability in a decade-long grassland experiment." *Nature* 441 (7093):629-632.

- Titman, George David 1976. "Interspecific competition for resources: an experimental and theoretical study." Ph. D. Thesis, The University of Michigan, Ann Arbor.
- Tokeshi, Mutsunori. 1990. "Niche apportionment or random assortment: species abundance patterns revisited." *The Journal of animal ecology*:1129-1146.
- Tokeshi, Mutsunori. 1993. "Species abundance patterns and community structure." *Advances in ecological research* 24:111-186.
- Tokeshi, Mutsunori. 1996. "Power fraction: a new explanation of relative abundance patterns in species-rich assemblages." *Oikos*:543-550.
- Ulrich, Werner. 2004. "Species co-occurrences and neutral models: reassessing JM Diamond's assembly rules." *Oikos* 107 (3):603-609.
- Ulrich, Werner, Marcin Ollik, and Karl Inne Ugland. 2010. "A meta-analysis of species-abundance distributions." *Oikos* 119 (7):1149-1155.
- Van Valen, Leigh M. 1985. "Null hypotheses and prediction." *Nature* 314 (6008):230-230.
- Vellend, Mark. 2010. "Conceptual synthesis in community ecology." *The Quarterly review of biology* 85 (2):183-206.
- Volkov, Igor, Jayanth R Banavar, Fangliang He, Stephen P Hubbell, and Amos Maritan. 2005. "Density dependence explains tree species abundance and diversity in tropical forests." *Nature* 438 (7068):658-661.
- Volkov, Igor, Jayanth R Banavar, Stephen P Hubbell, and Amos Maritan. 2003. "Neutral theory and relative species abundance in ecology." *Nature* 424 (6952):1035-1037.
- Volterra, Vito. 1926. "Fluctuations in the abundance of a species considered mathematically." *Nature* 118:558-560.
- Waters, C Kenneth. 1998. "Causal regularities in the biological world of contingent distributions." *Biology and philosophy* 13 (1):5-36.
- Watterson, Geoffrey A. 1974. "The sampling theory of selectively neutral alleles." *Advances in applied probability*:463-488.
- Weber, Marcel. 1999. "The aim and structure of ecological theory." *Philosophy of Science*:71-93.
- Whewell, William. 1847. *The philosophy of the inductive sciences, founded upon their history*. . 2 vols. Vol. 1. London: J.W. Parker.
- Whittaker, Robert H. 1965. "Dominance and Diversity in Land Plant Communities." *Science* 147 (3655):250-260.
- Wiens, John A. 1992. *The ecology of bird communities*. Vol. 1: Cambridge University Press.
- Williams, CB. 1947. "The generic relations of species in small ecological communities." *The Journal of Animal Ecology*:11-18.
- Williams, CB. 1964. *Patterns in the balance of nature*: Academic Press.
- Williams, Ernest E. 1969. "The ecology of colonization as seen in the zoogeography of anoline lizards on small islands." *Quarterly Review of Biology*:345-389.
- Wimsatt, William C. 1987. "False models as means to truer theories." In *Neutral models in biology*, edited by Matthew H Nitecki and Antoni Hoffman, 23-55. Oxford University Press.
- Worrall, John. 2003. "Normal science and dogmatism, paradigms and progress: Kuhn 'versus' Popper and Lakatos." In *Thomas Kuhn*, edited by Thomas Nickles, 65-100. New York: Cambridge University Press.
- Wright, S Joseph, and Stephen P Hubbell. 1983. "Stochastic extinction and reserve size: a focal species approach." *Oikos*:466-476.
- Zhou, Shu-Rong, and Da-Yong Zhang. 2008. "A nearly neutral model of biodiversity." *Ecology* 89 (1):248-258.