THE THEORY OF ECOLOGY Samuel M. Scheiner and Michael R. Willig, eds.

TABLE OF CONTENTS

INTRODUCTION

1. A General Theory of Ecology, Samuel M. Scheiner and Michael R. Willig

PERSPECTIVES ON THE ROLE OF THEORY IN ECOLOGY

- 2. Theory Makes Ecology Evolve, Jurek Kolasa
- 3. A General, Unifying Theory of Ecology?, Jay Odenbaugh

CONSTITUENT THEORIES OF ECOLOGY

- 4. Foraging Theory, Andrew Sih
- 5. Ecological Niche Theory, Jonathan Chase
- 6. Single Species Population Dynamics and its Theoretical Underpinnings, Alan Hastings
- 7. Natural Enemy-Victim Interactions: Do We Have a Unified Theory, Yet?, Robert Holt
- 8. The Metacommunity Concept and its Theoretical Underpinnings, Mathew A. Leibold
- 9. Domain and Propositions of Succession Theory, Steward T. A. Pickett, Scott J. Meiners, and Mary L. Cadenasso
- 10. The Equilibrium Theory of Island Biogeography, Dov Sax and Steven D. Gaines
- 11. Theory of Ecosystem Ecology, Ingrid C. Burke and William K. Lauenroth
- 12. Perspectives on Global Change Theory, Debra P. C. Peters, Brandon T. Bestelmeyer, and Alan K. Knapp
- 13. A Theory of Ecological Gradients, Gordon A. Fox, Samuel M. Scheiner, and Michael R. Willig
- 14. Biogeographical Gradient Theory, Robert K. Colwell

Synthesis

15. The State of Theory in Ecology, Michael R. Willig and Samuel M. Scheiner

Bibliography

LIST OF CONTRIBUTORS

Brandon T. Bestelmeyer USDA Agricultural Research Service Jornada Experimental Range

Ingrid C. Burke Department of Botany University of Wyoming

Mary L. Cadenasso Department of Plant Sciences University of California–Davis

Jonathan M. Chase Department of Biology Washington University

Robert Colwell Department of Ecology and Evolutionary Biology University of Connecticut

Gordon A. Fox Department of Biology University of South Florida

Steven D. Gaines Department of Ecology, Evolution and Marine Biology University of California–Santa Barbara

Alan Hastings Department of Environmental Science and Policy University of California–Davis

Robert Holt Department of Biology University of Florida

Alan K. Knapp Department of Biology Colorado State University William K. Lauenroth Department of Botany University of Wyoming

Mathew A. Leibold Section of Integrative Biology University of Texas–Austin

Scott J. Meiners Department of Biological Sciences Eastern Illinois University

Jay Odenbaugh Department of Philosophy Lewis & Clark College

Debra P. Peters USDA Agricultural Research Service Jornada Experimental Range

Stewart Pickett Institute for Ecosystem Studies

Dov Sax Department of Ecology and Evolutionary Biology Brown University

Samuel M. Scheiner Arlington, VA

Andrew Sih Department of Environmental Science and Policy University of California–Davis

Michael R. Willig Center for Environmental Sciences and Engineering University of Connecticut

Chapter 1: A General Theory of Ecology Samuel M. Scheiner and Michael R. Willig

"In the absence of agreed protocols and overarching theory, Ecology with its numerous subdisciplines, can sometimes resemble an amorphous, postmodern hotel or rabbit warren with separate entrances, corridors and rooms that safely accommodate the irreconcilable." (Grime 2007)

The development of theory in ecology is a lively and robust enterprise (Pickett et al. 2007). Despite claims to the contrary, the science of ecology has a long history of building theories that fruitfully guide research and deepen understanding. Our goal with this book is to reveal a selection of those theoretical structures. In doing so, our hope is that ecologists will better appreciate the theoretical frameworks within which they do research and will engage those theories in designing observational, experimental, and modeling components of their research. Many theories in ecology contain unspoken or even subconscious assumptions. By bringing such assumptions to the forefront, we can understand their consequences and discover new mechanisms or patterns, as well as new linkages among theories. Theory sometimes seems to be distant or disconnected from everyday practice in ecology. By the end of this book, the relevance of theory to understanding in ecology and its role in advancing science should become clear.

In this chapter, we present a general theory of ecology that serves as the supporting framework – a conceptual infrastructure – for the constitutive theories that appear in subsequent chapters. Chapters 2 (Kolasa) and 3 (Odenbaugh) consider the role that theory has played in ecology from the perspectives of a practicing ecologist and of a philosopher of science. The eleven chapters that make up the heart of the book delve into the theoretical underpinnings of a broad range of ecological subdisciplines. Although those chapters span the disciplinary range of ecology, they are representative rather than comprehensive. We could not possibly synthesize the full richness of ecological theory in a single book without it becoming encyclopedic. We encourage others to continue the process of theory development in other venues, and to reengage theoretical discourse with ecological research (e.g., Pickett et al. 2007).

We do not claim novelty for the general theory of ecology put forward here. Quite the contrary, the elements of the general theory have existed for at least the past 50 years. Many of

its principles are implicit in the tables of contents of most ecology textbooks, although our previous treatise (Scheiner and Willig 2008) was their first formal explication. In this chapter, we expand our earlier discussion of the structure of theories and the framework that underlies theory in ecology, providing a foundation for the chapters that follow in this book.

Importantly, we do not claim that the theory presented here is a final version. Rather, it should be considered provisional and ever changing, a general characteristic of theory that is often misunderstood by non-scientists. Indeed, the list of fundamental principles that we present will require additions, deletions, or refinements as ecological theory matures and is confronted by empirical evidence. Critically, this debate can occur only after the theory has been explicated. In the process of assembling this volume, we convened a workshop of the contributors at Center for Environmental Science and Engineering of the University of Connecticut. At that workshop, a fundamental principle emerged that was not considered in our previous paper (number 3 below).

THE STRUCTURE OF THEORIES

Before we present a general theory of ecology, we must describe the essence of theory and its structure (Tables 1.1 and 1.2). Theories are hierarchical frameworks that connect broad general principles to highly specific models. For heuristic purposes, we present this hierarchy as having three tiers (a general theory, constitutive theories, and models); however, we do not suggest that all theories fit neatly into one of these three categories. Rather, the framework will often stretch continuously from the general to the specific. The three tiers illustrate that continuum, and provide a useful way of viewing that hierarchy. The definitions and principles of the general theory are meant to encompass a wide variety of more specific constitutive theories, that in turn contain families of models. This view of constitutive theories as families of models is consistent with how theories are treated across all of biology and other sciences (van Fraassen 1980; Giere 1988; Beatty 1997; Longino 2002; Pickett et al. 2007; Wimsatt 2007; del Rio 2008; National Research Council 2008).

Each theory or model applies to a domain. The domain defines the universe of discourse – the scope of the theory – delimiting the boundaries within which constituent theories may be interconnected to form coherent entities. Constitutive theories are often most fruitful when they focus on one or a few factors in need of explanation (e.g., Hastings Chapter 6; Sax and Gaines

Chapter 10). Without such boundaries, we would be faced with continually trying to create a theory of everything.

Nonetheless, we recognize that domains are conceptual constructs and that theories or models may have overlapping domains. Changing the domain of a model can be a fruitful avenue for juxtaposing phenomena or processes that had been considered in isolation. For example, microeconomic theory uses three concepts – utility, income, and price – to understand consumer choices (Henderson and Quandt 1971; Mansfield 1979). Choices are assumed to maximize utility subject to income and price constraints. Behavioral ecologists essentially study the economics of choice for non-human animals, and have applied conceptual constructs and mathematical models from economics to understanding foraging ecology and space utilization (Stevens and Krebs 1986; see Sih Chapter 4). Recent examples of such borrowing of models across domains include the use of maximum entropy from thermodynamics theory (Harte et al. 2008; McRae et al. 2008) and connectivity models from electrical circuit theory (McRae et al. 2008).

All theories and models contain assumptions that simplify those models and place the focus on the explicit variables. The problem with many assumptions is that they are unstated, even subconscious in nature. Making such assumptions explicit sometimes may change the focus of the theory. For example, a fundamental principle of ecology is that ecological traits arise through evolution, but nearly always this is an unstated and oft ignored assumption. Models of community assembly usually ignore phylogenetic relationships among species. Recently, models that incorporate phylogenetic relationships have added substantially to our understanding of community assembly (e.g., Kraft et al. 2007).

Sometimes, such unstated assumptions can turn around and bite us. Most models of life history evolution assume that organisms can always adopt the optimal phenotype, instantaneously reallocating resources from growth to reproduction, and so ignoring evolutionary and developmental constraints. Ignoring this assumption led to predictions that were biologically improbable, e.g., an organism should allocate 100% of its resources to reproduction one day after it devoted 100% of its resources to growth (Schaffer 1983), or an annual plant should switch multiple times between growth and reproduction (King and Roughgarden 1982).

Principles and propositions

When asked to describe a theory, we often think about a set of broad statements about empirical patterns and the processes that operate within a domain. For the sake of clarity, we use different terms to refer to those broad statements when we speak of general theories (fundamental principles) versus when we speak of constitutive theories (propositions). In part, fundamental principles are similar to propositions. Each can be concepts (labeled regularities) or confirmed generalizations (condensations of facts). They differ in that fundamental principles are broader in scope, often encompassing multiple, interrelated patterns and mechanisms. Because constitutive theories are meant to guide the building of specific models, their propositions should be more precise statements that represent the potential individual components of those models.

Propositions can consist of laws: statements of relationship or causation. The propositions are where the fundamental principles of the general theory are integrated. For the general theory of ecology, some of the principles involve patterns, others involve processes, many involve both (see below). Thus, the causal linking of process and pattern, the law-like behavior that we look for in theories, occurs through the propositions of the constitutive theories.

Laws reside within constitutive theories, and not as part of the general theory, because no law is required for the construction of models in all of ecology's subdomains. Several chapters show, however, that ecology is rich in laws that hold within more limited domains (see discussion in Willig and Scheiner Chapter 15). A brisk debate has occurred over whether ecology has any laws at the level of the general theory (e.g., Lawton 1999; Murray 2000; Turchin 2001; Berryman 2003; Simberloff 2004; O'Hara 2005; Pickett et al. 2007; Lockwood 2008), which is related to the debate about laws across all of biology (e.g., Beatty 1997; Brandon 1997; Mitchell 1997; Sober 1997). The continuing search for such laws is an important aspect of a theory's evolution.

The reaction of many to confirmed generalizations is, "Well, isn't that obvious?" In reality, the answer is no. Often such generalizations are obvious only after their explication. Generalizations serve as reminders about assumptions contained in lower level theories or models. For example, a fundamental principle in ecology is that ecological processes depend on contingencies (see below). Yet, many ecological theories and models are deterministic, and ignore the role of contingency or stochasticity in molding patterns and processes in nature. Deterministic models are not wrong, just potentially incomplete. Sometimes ignoring contingencies has no effect on model prediction. At other times, the consequences can be

profound. As the statistician George E. P. Box is reputed to have said, "Essentially, all models are wrong, but some are useful."

Fundamental principles keep prodding us to test assumptions. For example, one fundamental principle tells us that species are comprised of individuals that differ in phenotype. Nonetheless, many ecological theories assume that species consist of identical individuals. Although this is a useful simplification in many instances, it is important to be reminded continually about this assumption and its consequences to predictive understanding. Similarly, many of the fundamental principles consider variation in the environment or species interactions, yet many constitutive theories or models average over that variation.

Not all assumptions within a constitutive theory derive from the fundamental principles of the general theory. Some assumptions derive from other domains. If an assumption is taken unchanged from another domain it may be unspecified within a theory. For example, all constitutive theories in ecology take as given the conservation of matter and energy, fundamental principles from the domain of physics. We take as given the fundamental principles of any other general theory. As such, we recognize the general tenet of consilience: the entire set of scientific theories must be consistent with each other (Whewell 1858). The decision to explicitly include such assumptions as fundamental principles within the theory under consideration depends on whether those assumptions are subject to test within that theory. Since no theory in ecology would ever test the conservation of matter, it lies outside the theory.

Theories may clash, but such clashes indicate foci of research that advance understanding. In general, theories inhabiting different domains will not clash directly, although results from one domain can point to problems with theories in other domains. For example, studies of geographical distributions of clades of organisms within the domain of historical biogeography became important evidence for the theory of continental drift, a part of the domain of geology. In that instance, the need for a causal mechanism to explain distribution patterns, was a factor that led to the development of new fundamental principles in another domain.

Models

At the lowest level of our theory hierarchy are models. Models are where the theoretical rubber meets the empirical road. Many ecological theories are just such models. Although scientific theories encompass a wide variety of types of models, including physical models (e.g., Watson

and Crick's ball and wire model of a DNA molecule), in ecology we generally deal with abstract or conceptual models. These models may be analytic, statistical, or simulations.

Recognizing that what is often labeled as a theory is but one model within a larger theory can help to clarify our thinking. For example, Scheiner and Willig (2005) assembled an apparently bewildering array of 17 models about species richness gradients into a framework built on just four propositions. A similar process of clarification can be found in Chapters 8, where Leibold shows that all metacommunity theories can be captured within a single framework of just two variables: amount of interpatch heterogeneity and dispersal rate. Other chapters in this book provide further examples of model unification.

Because theories often consist of families of models, it is possible for models to be inconsistent or even contradictory. Sometimes, such inconsistencies point to areas that require additional empirical evaluation or model development. But sometimes contradictory models can be maintained side-by-side because they serve different functions or are useful under different conditions. For example, in some physics models, light is treated as a particle and in others as a wave. There is thus no need to insist that contradictory models always be reconciled within themselves or that one need always prevail. Instead, this apparent contradiction is resolved at a higher level in the theory hierarchy by a more general theory that allows for both wave-like and particle-like behavior of light. The apparently contradictory models are built from differing sets of propositions arising from different assumptions and, thus, refer to different domains. In a similar fashion, constitutive theories can be contradictory if they are built with different assumptions.

THE DOMAIN OF ECOLOGY

The domain of ecology is the spatial and temporal patterns of the distribution and abundance of organisms, including causes and consequences. Although our definition of the domain spans the definitions found in most textbooks (Pickett et al. 2007; Scheiner and Willig 2008), it is different in two respects. First, our definition includes the phenomena to be understood (i.e., spatial and temporal patterns of abundance of organisms) and the causes of those phenomena. Some definitions include only the latter (i.e., interactions of organisms and environments). Second, and most strikingly, our definition explicitly includes the study of the consequences of those phenomena, such as the flux of matter and energy.

In general, the domain of a theory defines the objects of interest and their characteristics. Ecological theories make predictions about three types of objects: species, individuals, and traits or consequences of individuals. Parts of ecology (e.g., ecosystem theory) also makes predictions about fluxes and pools of elements and energy. However, what makes these theories part of the domain of ecology is that those fluxes and pools are controlled or affected by organisms. Thus, they are aggregate consequences of species, individuals, or the traits of individuals. Otherwise, ecosystem theory would reside firmly in the domain of the geosciences.

All three types of objects share an important property, variability (see below). This collection of objects distinguishes ecology from other, related and overlapping domains. The theory of evolution makes predictions about species and the traits of individuals. Its domain differs from that of ecology in that predictions are always about collections of individuals (e.g., gene frequencies), never about a single individual. In contrast, theories in ecology may make predictions about either collections of individuals or a single individual (e.g., Sih Chapter 4). Because a given object may be part of multiple domains, understanding of that object and its characteristics depends on examining it within the context of all of those domains.

Just as a general theory has a domain, each constitutive theory and particular model has a domain. Explicitly defining each such domain is important for two reasons. First, a domain defines the most central or general topics under investigation. Second, a clear definition indicates which objects or phenomena are excluded from consideration. Many protracted debates in ecology have occurred when proponents or opponents of particular theories or models have attempted to make claims that fall outside a theory's domain. For example, the extensive debates over the causes of large-scale patterns of plant diversity (e.g., Huston 1994; Waide et al. 1999; Mittelbach et al. 2001; Mittelbach et al. 2003; Whittaker and Heegaard 2003) are based on extrapolating to continental and global scales, models that are valid only at a regional scale (Fox et al. Chapter 13).

OVERLAPPING DOMAINS

The domain of the theory of ecology overlaps substantially with several other domains. Of course, all scientific domains overlap in some fashion, but we speak here of those domains that make predictions about some of the same objects of study as does the theory of ecology, or constitutive theories that use fundamental principles from other domains. A constitutive theory

can straddle two or more general theories if some of its models ultimately address a central question of each general theory. One way to decide if a constitutive theory straddles two general theories is to consider the assumptions of those general theories. If the constitutive theory simply accepts all of the assumptions in a particular general theory and never questions or tests them, it likely is not a member of that general theory.

A corollary of the previous statement is that any given model of necessity explores or tests one or more of the assumptions, fundamental principles or propositions of a theory. For example, a continuing issue in ecology concerns the identity of parameters that can be treated as constants and those that need to be treated as variables in a particular theory or model. If a parameter is treated as a constant, the average value of that parameter is assumed to be sufficient because either the variation has no effect or acts in a strictly additive fashion relative to the causative mechanisms under examination.

In some instances, ecologists make assumptions without ever testing them. For example, it is reasonable to assume that we can average over quantum fluctuations (from the domain of physics) in ecological processes. On the other hand, the physiological variations that occur from minute to minute in a mammal so as to maintain body temperature (from the domain of the theory of organisms) (Scheiner submitted; Zamer and Scheiner in prep.) may matter for ecological processes and should not be averaged in some instances. For example, basal metabolic rates in large mammals can vary substantially between winter and summer. Failure to account for this variation can seriously overestimate winter energy expenditures and food intake requirements (Arnold et al. 2006).

A subdomain can overlap two domains. For example, ecosystem science has some constituent theories that are part of ecology and some that are part of geosciences. Such overlaps can extend to the level of individual models. For example, foraging theory (Sih Chapter 4) contains some models that are ecological, others that are evolutionary, and others that are both. This sharing of subdomains shows that the boundaries of domains are not distinct and can be somewhat arbitrary.

A domain as defined by a general theory, constitutive theory, or model should be a coherent entity. Some named areas are not domains, but collections of domains. For example, evolutionary ecology consists of a set of constituent theories, some of which are within the domain of the theory of ecology and others that are within the domain of the theory of evolution.

Physiological ecology is likely another such collection of domains, although at this time it is difficult to tell because a theory of physiology has not yet been articulated.

THE FUNDAMENTAL PRINCIPLES OF ECOLOGY

The general theory of ecology consists of eight fundamental principles (Table 1.3). The roots of these principles can be traced to the origins of ecology in the 19th century. They were in place and widely accepted by the 1950s, were recently codified as the components of a general theory (Scheiner and Willig 2008), and continue to evolve. In particular, we have added an eighth fundamental principle (number 3) so that the numbering of this set differs somewhat from our previous list.

Heterogeneous distributions

The first fundamental principle – the heterogeneous distribution of organisms – is a refinement of the domain of the theory of ecology. The heterogeneity of distributions is one of the most striking features of nature: all species have a heterogeneous distribution at some if not most spatial scales. Thus, this principle encompasses a basic object of interest, is its most important property, and serves to guide the rest of the theory. All of the other parts of the theory of ecology serve to either explain this central observation or to explore its consequences. Arguably, the origins of ecology as a discipline and the first ecological theories can be traced to its recognition (Forster 1778; von Humboldt 1808). This heterogeneous distribution is both caused by and a cause of other ecological patterns and processes.

Environmental interactions

The second fundamental principle – interactions of organisms – includes within it the vast majority of ecological processes responsible for heterogeneity in time and space. They include both intraspecific and interspecific interactions such as competition, predation, and mutualism, as well as feedbacks between biotic and abiotic components. Within this principle, particular interactions that are part of constituent theories act to elaborate the general theory (see later chapters). Many definitions of ecology are restatements of this principle (Scheiner and Willig 2008).

Variation of individuals

The third principle – the variation of individuals – is the result of processes that derive from the theory of organisms (Scheiner submitted; Zamer and Scheiner in prep.). Ecological theories make predictions about the characteristics or aggregate properties of species, individuals or traits. The majority of ecological theories make predictions about species or collections of species (e.g., species richness of communities; see Chapters 8-10, 13, 14). Some theories, such as population ecology and behavioral ecology, concern themselves with predictions about individuals or collections of individuals (e.g., numbers of individuals in a population; see Chapters 4-8). Some theories make predictions about the properties of individuals or species (e.g., body size distributions; see Chapters 4, 5, 8, 10, 13, 14). Finally, some theories make predictions about the aggregate properties of individuals or species (e.g., ecosystem standing biomass; see Chapter 11).

Groups of species or individuals share the property that the members of those groups differ in their characteristics, even though many theories and models assume invariance. For example, one of the most common hidden assumption in models of species richness is that all individuals within a species are identical (e.g., Fox et al. Chapter 13). Such assumptions may be reasonable for the purposes of simplifying models. Violations of this assumption may not substantially change predictions. However, in some cases relaxing this assumption has led to substantial changes in predictions. For example, when the chances of survival are allowed to vary among individuals within a population, treating all individuals as equal turns out to substantially misestimate the risk of local extinction from demographic stochasticity; depending on the model used for reproduction, treating all individuals as identical can over- or underestimate that risk (Kendall and Fox 2003).

Contingency

The fourth fundamental principle – contingency – has grown in importance in ecological theory and now appears in a wide variety of constituent theories and models. By contingency we mean the combined effects of two processes – randomness and sensitivity to initial conditions. Contingency is an important cause of the heterogeneous distribution of organisms, both at very large and very small extents of time and space (e.g., a seed lands in one spot and not another; a particular species arises on a particular continent). This principle exemplifies the dynamic nature

of a theory. A theory is constantly evolving, although substantive change typically occurs over decades. One hallmark of that dynamic is the emergence of new principles, such as occurred with this principle during the 1960s to 1980s.

Heterogeneity of environmental conditions

The fifth fundamental principle – environmental heterogeneity – is a consequence of processes from the theories of earth and space sciences when the environmental factors are abiotic, as well as the consequences of the second principle when those factors are biotic. For example, seasonal variation in temperature is the result of orbital properties of the Earth, whereas a variety of geophysical processes create heterogeneity in environmental stressors like salt (e.g., wave action near shores) or heavy metals (e.g., geologic processes that create differences in bedrocks). This principle is part of many constituent theories and contains a broad class of underlying mechanisms for the heterogeneous distribution of organisms, as seen in many of the constitutive theories presented in this book. As with the second principle, particular mechanisms pertain to particular constituent theories.

Finite and heterogeneous resources

The sixth principle – finite and heterogeneous resources – is again a consequence of processes from the theories of earth and space sciences or the second principle. Although variation in resources is similar to variation in environmental conditions, a fundamental distinction is the finite, and thus limiting, nature of these resources. Unlike an environmental condition, a resource is subject to competition. For example, seasonal variation in light and temperature are caused by the same orbital mechanisms, but light is subject to competition (e.g., one plant shades another) whereas temperature is a condition and not subject to competition. This distinction in the nature of environmental factors with regard to competitive processes can result in different ecological outcomes. For example, β diversity in plant communities is high in warm deserts and low in arctic tundra because diversity in warm deserts is controlled by water, a limiting resources, while diversity in arctic tundra is controlled by temperature, an environmental condition (Scheiner and Rey-Benayas 1994). Whether a particular environmental factor is a condition or a resource can be context dependent. For example, water is sometimes a resource subject to competition (e.g., plants in a desert) and sometimes a condition (e.g., fish in the ocean). Some heavy metals (e.g., manganese) can be limiting to plants if at low levels, so acting as a resource, and be toxic at high levels, so acting as a condition.

Birth and death

The seventh fundamental principle – the birth and death of organisms – is the result of processes that come from the domain of organismal biology: physiology and development (Scheiner submitted; Zamer and Scheiner in prep.). One of the fundamental characteristics of life is reproduction. While birth comes about through cellular and organismal processes, such as fertilization and development, the rate that it occurs depends on interactions of an organism with its environment, such as the uptake of nutrients or mating.

Similarly, a defining characteristic of life is that all organisms are mortal. By "mortal" we mean that no organism is invulnerable, i.e., any organism might die as the result of predation, stress, trauma or starvation. Thus, the rate of death depends on environmental interactions. We do not mean that all organisms senesce. The senescence of organisms, a decrease in function or fitness with age, is a more narrow version of this principle that would apply to particular constituent theories. This fifth principle forms the basis of a large number of constituent theories concerning phenomena as wide ranging as life histories, behavior, demography, and succession (e.g., Chapters 4, 6 and 9).

Evolution

The eighth principle – the evolutionary cause of ecological properties – is the result of processes that derive from the theory of evolution. The inclusion of evolution within ecological thinking was an important outcome of the Modern Synthesis. Although evolutionary thinking about ecological processes goes back at least to Darwin (1859), evolutionary thinking had been influencing ecology widely since at least the 1920s (Collins 1986; Mitman 1992) and its widespread acceptance occurred primarily in the latter half of the 20th century. The acceptance of this principle led to such disciplines as behavioral ecology (Sih Chapter 4) and population biology, and contributed to the demise of the Clementsian superorganism theory.

This principle illustrates how theories in overlapping domains can interact with each other. One of the fundamental principles of the theory of evolution is that evolutionary change is caused primarily by natural selection (Mayr 1982). Fitness differences among individuals, a key

component of the process of natural selection, is caused in large part by ecological processes. So, ecology drives evolution which, in turn, determines ecological properties.

15

OVERVIEW

This chapter only begins to delve into the many issues relating to theory structure and development in ecology. For a much more comprehensive discussion, we recommend Pickett *et al.* (2007). One purpose in articulating a general theory is to clarify thinking, and bring to the fore aspects of science that may not be recognized consciously. For example, it is notable that five of the eight fundamental principles are about variability. Although ecologists sometimes decry the variation among the entities that they study and claim that such variation prevents the development of laws or predictions, we suggest that progress in ecology requires that ecologists embrace this variation and explicitly encompass it in theories. More important, recognizing that variation is a pervasive property of our discipline helps explain why ecologists sometimes have difficulty communicating about ecology to colleagues in other disciplines, where the focus is on the shared properties of organisms rather than on their variability.

Throughout the process of developing and articulating the general theory and the constitutive theories of ecology, we have been impressed by how often the statement and full consideration of the obvious can lead to deep insights. The chapters that follow demonstrate that process. Our hope is that such insights will substantially improve how we do our science. Ecologists often despair over the seemingly endless variety of their science with no clear overarching structure. The theories discussed in this book present a critical set of steps in unifying that structure.

Acknowledgements

We thank Todd Crowl, Jay Odenbaugh, and Steward Pickett for thoughtful comments on an earlier draft. Many of the ideas presented in this chapter emerged from or were clarified by a workshop of the contributors to this book. We thank all of the participants for their stimulating interactions. Support to MRW was provided by the Center for Environmental Sciences and Engineering at the University of Connecticut. Support for a workshop that brought the chapter authors together was provided by the Center for Environmental Sciences and Engineering in cooperation with the Office of the Vice Provost for Research and Graduate Education at the University of Connecticut. The views expressed in this paper do not necessarily reflect those of the National Science Foundation or the United States Government.

Literature cited

- Arnold, W., T. Ruf, and R. Kuntz. 2006. Seasonal adjustment of energy budget in a large wild mammal, the Przewalski horse (*Equus ferus przewalskii*) II. Energy expenditure. Journal of Experimental Biology 209:4566-4573.
- Beatty, J. 1997. Why do biologists argue like they do? Philosophy of Science 64:S432-S443.
- Berryman, A. A. 2003. On principles, laws and theory in population ecology. Oikos 103:695-701.
- Brandon, R. N. 1997. Does biology have laws? The experimental evidence. Philosophy of Science 64:S444-S457.
- Collins, J. P. 1986. Evolutionary ecology and the use of natural selection in ecological theory. Journal of the History of Biology 19:257-288.
- Darwin, C. R. 1859, On the Origin of Species by Means of Natural Selection. London, UK, Murray.
- del Rio, C. M. 2008. Metabolic theory or metabolic models? Trends in Ecology & Evolution 23:256-260.
- Forster, J. R. 1778, Observations made during a voyage round the world, on physical geography, natural history, and ethnic philosophy. London, UK, G. Robinson.
- Giere, R. N. 1988, Explaining Science: A Cognitive Approach Chicago, IL, University of Chicago Press.
- Grime, J. P. 2007. Plant strategy theories: a comment on Craine (2005). Journal of Ecology 95:227-230.
- Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. Maximum entropy and the state-variable approach to macroecology. Ecology 89:2700-2711.
- Henderson, J. M., and R. E. Quandt. 1971, Microeconomic Theory: A Mathematical Approach. New York, New York, McGraw Hill.
- Huston, M. A. 1994, Biological Diversity: the coexistence of species in changing landscapes. Cambridge, UK, Cambridge University Press.
- Kendall, B. E., and G. A. Fox. 2003. Unstructured individual variation and demographic stochasticity. Conservation Biology 17:1170-1172.
- King, D., and J. Roughgarden. 1982. Multiple switches between vegetative and reproductive growth in annual plants. Theoretical Population Biology 21:194-204.

- Kraft, N. J. B., W. K. Cornwell, C. O. Webb, and D. D. Ackerly. 2007. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. American Naturalist 170:271-283.
- Lawton, J. H. 1999. Are there general laws in ecology? Oikos 84:177-192.
- Lockwood, D. R. 2008. When logic fails ecology. Quarterly Review of Biology 83:57-64.
- Longino, H. E. 2002, The Fate of Knowledge. Princeton, NJ, Princeton University Press.
- Mansfield, E. 1979, Microeconomics: Theory and Applications. New York, New York, W.W. Norton.
- Mayr, E. 1982, The Growth of Biological Thought. Cambridge, MA, Belknap Press.
- McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712-2724.
- Mitchell, S. D. 1997. Pragmatic laws. Philosophy of Science 64:S468-S479.
- Mitman, G. 1992, The State of Nature. Chicago, IL, University of Chicago Press.
- Mittelbach, G. G., S. M. Scheiner, and C. F. Steiner. 2003. What is the observed relationship between species richness and productivity? Reply. Ecology 84:3390-3395.
- Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig et al. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381-2396.
- Murray, B. G., Jr. 2000. Universal laws and predictive theory in ecology and evolution. Oikos 89:403-408.
- National Research Council. 2008, The Role of Theory in Advancing 21st Century Biology: Catalyzing Transformative Research. Committee on Defining and Advancing the Conceptual Basis of Biological Sciences in the 21st Century. Washington, DC, National Academies Press.
- O'Hara, R. B. 2005. The anarchist's guide to ecological theory. Or, we don't need no stinkin' laws. Oikos 110:390-393.
- Pickett, S. T. A., J. Kolasa, and C. G. Jones. 2007, Ecological Understanding: The Nature of Theory and the Theory of Nature. New York, NY, Elsevier.
- Schaffer, W. N. 1983. On the application of optimal control theory to the general life history problem. American Naturalist 121:418-431.

- Scheiner, S. M. submitted. Towards a conceptual framework for biology. Quarterly Review of Biology.
- Scheiner, S. M., and J. M. Rey-Benayas. 1994. Global patterns of plant diversity. Evolutionary Ecology 8:331-347.
- Scheiner, S. M., and M. R. Willig. 2005. Developing unified theories in ecology as exemplified with diversity gradients. American Naturalist 166:458-469.
- —. 2008. A general theory of ecology. Theoretical Ecology 1:21-28.
- Simberloff, D. 2004. Community ecology: is it time to move on? American Naturalist 163:787-799.
- Sober, E. 1997. Two outbreaks of lawlessness in recent philosophy of biology. Philosophy of Science 64:S458-S467.
- Stevens, D. W., and J. R. Krebs. 1986, Foraging Theory. Princeton, NJ, Princeton University Press.
- Turchin, P. 2001. Does population ecology have general laws? Oikos 94:17-26.
- van Fraassen, B. 1980, The Scientific Image. Oxford, UK, Oxford University Press.
- von Humboldt, A. 1808, Ansichten der Natur mit wissenschaftlichen Erlauterungen. Tübingen, Germany, J. G. Cotta.
- Waide, R. B., M. R. Willig, C. F. Steiner, G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday et al. 1999. The relationship between productivity and species richness. Annual Review of Ecology and Systematics 30:257-300.
- Whewell, W. 1858, Novum Organon Renovatum. London, UK.
- Whittaker, R. J., and E. Heegaard. 2003. What is the observed relationship between species richness and productivity? Comment. Ecology 84:3384-3390.
- Wimsatt, W. C. 2007, Re-engineering Philosophy for Limited Beings. Cambridge, MA, Harvard University Press.

Table 1.1. A hierarchical structure of theories including it components. A general theory creates the framework within which constitutive theories can be articulated, which in turn sets the rules for building models. Conversely, tests of models may challenge the propositions and assumptions of its constitutive theory, which in turn may result in a change in the fundamental principles of the general theory. See Table 2 for definitions of terms.

General Theory

Background: domain, assumptions, framework, definitions Fundamental principles: concepts, confirmed generalizations Outputs: constitutive theories

Constitutive Theory

Background: domain, assumptions, framework, definitions Propositions: concepts, confirmed generalizations, laws Outputs: models

Model

Background: domain, assumptions, framework, definitions, propositions

Construction: translation modes

Outputs: hypotheses

Tests: facts

Component Description Assumptions Conditions or structures needed to build a theory or model Concepts Labeled regularities in phenomena Confirmed generalizations Condensations and abstractions from a body of facts that have been tested Definitions Conventions and prescriptions necessary for a theory or model to work with clarity Domain The scope in space, time, and phenomena addressed by a theory or model Facts Confirmable records of phenomena Nested causal or logical structure of a theory or model Framework Testable statements derived from or representing various Hypotheses components of the theory or model Laws Conditional statements of relationship or causation, or statements of process that hold within a universe of discourse Model Conceptual construct that represents or simplifies the natural world Translation modes Procedures and concepts needed to move from the abstractions of a theory to the specifics of model, application or test

Table 1.2. Definitions of terms for the theory components in Table 1.1. (modified from Pickett et al. 2007)

Table 1.3. Eight fundamental principles of the general theory of ecology (modified from Scheiner and Willig 2008)

- 1. Organisms are distributed in space and time in a heterogeneous manner.
- 2. Organisms interact with their abiotic and biotic environments.
- 3. Variation in the characteristics of organisms results in heterogeneity of ecological patterns and processes.
- 4. The distributions of organisms and their interactions depend on contingencies.
- 5. Environmental conditions are heterogeneous in space and time.
- 6. Resources are finite and heterogeneous in space and time.
- 7. Birth rates and death rates are a consequence of interactions with the abiotic and biotic environment.
- 8. The ecological properties of species are the result of evolution.

Chapter 2: Theory Makes Ecology Evolve Jurek Kolasa

He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast.

-Leonardo da Vinci (1452-1519)

In this chapter I will comment on the importance of theoretical thought in the development of ecology. I will pursue two themes. One draws on history of ecological theories or ideas that led to formulation of theoretical frameworks and another reviews benefits and limitations of using theory in the broad sense. Finally, using these two perspectives, I conclude with comments on the recent proposition to erect a general theory of ecology.

Determination of when particular theoretical ideas entered and influenced ecology relies on making judgments. However, evaluating the role of theory in ecology is not easy, particularly because some aspects of theory are not generally acknowledged as such. Indeed, the task is rather straightforward when we consider quantitative models (but see Scheiner and Willig, Chapter 1, and Odenbaugh, Chapter 3, for the distinction between models and theories; also Pickett et al. 2007) and propositions that are recognized by most if not all ecologists as belonging to the realm of theory. However, not all theory is recognized immediately as such, particularly when it is not quantitative or was not traditionally discussed as theory in ecological literature. Qualitative, sometimes informal theoretical ideas play an important role in focusing, refining, and advancing ecological research as well. To acknowledge and assess their role, it is necessary to present some broad criteria for distinguishing between purely empirical work, if indeed such a thing exists at all, and ideas of substantial theoretical content. Sometimes theoreticians recognize explicitly that math is preceded by crucial conceptualizations. Andrew Sih (Chapter 4) says, for example, that the key challenge for foraging theory (and for optimality theory, in general) is not the math but capturing the key elements of the biology of the system. Again, that task appears much easier in retrospect, but it is much harder when theoretical ideas are in the early stages of development.

Another difficulty in evaluating the role of theory stems from the low cohesion of concepts, generalizations, and mathematical models in relation to one another. The lack of cohesion contributes to the intellectual immaturity of ecology (Hagen 2003) in spite of ecology's

spectacular growth. Many of the observations identified as the principles of Scheiner and Willig (Table 1.3) have been recognized at various stages of development of ecology, but were not collected methodically into a system of propositions. Once they are, their theoretical significance becomes unambiguous. Whenever possible I will try to identify their first articulations.

Biologists were interested in ecological questions well before the science of ecology emerged as an identifiable discipline. Indeed, if one defined the domain of ecology as emergence and interactions of ecological entities (organisms and supra-organismal formations such as family groups, herds, or communities although emergence of organisms is an exception as it falls into the domain of developmental biology) with each other and with their environment (but see Scheiner and Willig 2008, also Chapter 1), many biological processes would be included. Any consideration of surrounding biological phenomena is likely to involve some ecology. Greek philosophers were interested in things that today fall within the scope of ecology. Theophrastus developed the conception of an autonomous nature, interacting with man and described various interrelationships among animals and between animals and their environment as early as the 4th century BC (Ramalay, 1940). His accounts entitled "On the Causes of Plants) represent the first known efforts to organize, interpret, and expand knowledge of plant reproduction, requirements, and uses. These early interests in the relations between an traits of organisms and the environment grew slowly. Initial questions were pragmatic and with limited theoretical content. Naturalists of the 18th and 19th centuries frequently had interest in ecological interactions, even if such interactions were not their main focus. Darwin was acutely aware of the importance of competition and predation, and made a spectacularly successful use of these notions in developing the assumptions of evolutionary theory.

Systematic exploration of such pragmatic questions began in earnest in England. The Rothamsted Experimental Station, founded in 1843, initiated a series of experiments between 1843 and 1856 that were aimed at a variety of applied, yet clearly ecological problems related to agriculture.

Parallel to the practical concerns, ideas that we recognize as modern began to emerge as well. The ecological concept of integrated communities of organisms first appeared in the 19th century with the studies of A. Grisebach (1838), a German botanist. The importance of this perspective was reflected in the need to define the discipline. In 1866, Haeckel recognized and

captured the multifaceted complexity of interrelationships among organisms (cf. Figure 2.1). He coined the term ecology to carve a separate scope for investigation of those relationships. Other biologists followed Grisbach's ecological approach to natural history studies: Möbius (1877) investigated Danish oyster banks whereas Forbes (1887) described a lake community as a microcosm (a conceptualization that laid the foundation for important modern views such as the ecosystem concept or habitat-based conservation). Many others subsequently explored the ecology of water bodies, agricultural systems, or forests. The first professional journal in this area, The American Naturalist, started publishing in 1867.

Yet, when we are asked about the beginnings of ecology, we will, most likely, hear the names of Forbes, Cowles, Clements, Tansley, Lotka, Volterra, Gause, Gleason and others, all working in the early 20th century. The question one might pose is why these names specifically? Much was known about natural phenomena before these scientists gained the recognition they deserve. So, what made them famous while dozens of others like them, who investigated lakes or vegetation changes, are largely left out of modern accounts?

These individuals stand out in ecology textbooks not because they are associated with explorations of the natural world, but because they proposed new and broad ideas or because they presented empirical studies within the framework of such ideas. Sometimes, they offered formal and flexible tools for developing theory (e.g., Lotka, Volterra, Gause) while sometimes they framed broad classes of phenomena as powerful and appealing concepts (Forbes, Cowles, Clements, Gleason, and Tansley).

THEORETICAL COMPONENTS

Before we examine the role of theory in the development of ecology, it is useful to note that scientific theory is a broad notion. Its borders may be fuzzy. This is because any empirical research is permeated by theoretical constructs, and any theory has an empirical content, at least in the natural sciences. A couple of extreme examples illustrate this point. Most ecologists would view measurements of temperature as a purely empirical activity. I agree that collecting temperature data is an empirical activity. However, temperature, defined as a mean kinetic energy of molecules, is a complex theoretical construct. Mean is a theoretical term. Kinetic energy is a theoretical term. A molecule is a rich theoretical conceptualization of small portions of matter (also a theoretical notion). Finally, temperature indicators, whether mercury, bimetallic

sensor or some other device, involve a whole suite of theoretical assumptions that link temperature to directly observed behavior of some other material processes.

Thus, while measuring temperature data is an empirical activity, in doing so ecologists subscribe to a plethora of theoretical concepts. At the other end of the spectrum lie concepts such as a community. Community ecologists routinely measure community metrics, which implies acceptance of a number of theoretical premises such as the existence of links between the metrics and community attributes. For example, the rationale for measuring evenness or Shannon-Wiener diversity indices relies on the belief that they convey information about processes determining the distribution of abundances, which may or may not be true. Meaningful empirical work is difficult to imagine without first accepting a set of theoretical constructs.

Indeed, because empirical and theoretical constituents combine to form various terms that ecologists use, it is difficult to clearly separate the theoretical and observational components to the development of understanding. Such a difficulty may be particularly serious when the conceptual device is not fully developed (one might think, for example, of a historical situation when heat and temperature were not yet clearly defined as different concepts). In ecology, such difficulties arose when the potential for, process of, and the result of competition were used interchangeably to some degree and stimulated a debate about 'the ghost of competition past' (e.g., Strong, 1984, Keddy, 1989). Furthermore, some situations, such as interference competition, may fail to fit neatly either the process-based view of competition or the outcomebased view because it emphasizes one-on-one interactions among individuals as opposed to the population level outcome or population level resource limitations. Different meanings may lead to different experimental tests and different interpretations of results (Pickett et al. 2007).

The critical question we thus face is when we can usefully credit theory as a driving factor for scientific advances and when credit be given to empirical discoveries. I will not attempt to resolve this methodological question. Rather, I will restate ideas that may help in addressing the role of theory in shaping progress in ecology, without claiming that these ideas represent more than a convenience. Addressing this question requires accepting that theory and its various components exist at different levels of development (articulation, clarity, formalization). When ecologists explicitly propose a model such as a competition equation, we have no difficulty accepting it as theoretical construct. When however the theoretical facet of a

proposition is less formal, some might raise their brows. So, how do we decide that a proposition has significant theoretical contents? One way of answering this question is to examine a proposition for the nature of its central notions.

Notions

Early in the development of a theory, scientists may employ imprecise or incompletely articulated concepts, metaphors, or analogies – notions. Because notions do not have explicit assumptions or form consistent conceptual structure, they are not considered parts of theory per se. Nevertheless, even at an early stage of development, they may influence the growth of theory as well as inspire and stimulate empirical research. When applicable and convenient, I will identify notions that were seminal to advancement of understanding.

Some of the early proposed notions developed into more formalized structures (e.g., niche), others continue to function informally (e.g., source-sink dynamics or even metacommunity). Furthermore, current definitions of metacommunity (e.g., Leibold Chapter 8) rely on a vague criterion of local communities being "linked by dispersal". This criterion works well at this stage of theory development in spite of being rather unruly: We do not know how much dispersal is needed, at what time scales, and whether it includes all, some, or just one species (M. Leibold feels one is sufficient, though). Depending on the arbitrary decisions regarding such specifications, a metacommunity either comprises just one community, the whole earth, or anything in between. Clearly, notions can be helpful without being conceptually polished. Still some notions have been recognized as confusing and therefore not very useful because the confusion may hinder progress (e.g., balance of nature, cf. Wiens 1984).

Generalizations and idealizations

Two other terms are much more diagnostic of theory and thus easier to interpret in the course of ecology growth and development. Both have roots in the inductive process. Without delving into the fine distinctions between them, I will attempt to identify generalizations and idealizations that have been made at different times by ecologists, particularly when they led to a flurry of conceptual or empirical activity. In brief, generalization is a statement reflecting regularities extracted from a set of observations – in common parlance is it a condensation of facts. Although a generalization, by virtue of the process through which it is arrived at, has

substantial empirical content, it acquires theoretical flavor during its articulation. It is then that its scope is specified (which observations qualify for inclusion), relationships are exposed (what is shared among these observations), and language (definitions of facts and relationships), sometimes formal, settled. For example, in this book (Scheiner and Willig Chapter 1), the generalization that individuals are different (i.e., share the property of being different) from each other is clearly based on numerous observations. However, it is not a sum of observations by any means. The generalization uses a relationship between two individuals of "being different". Thus, a condensation of observations is produced by introducing a new concept, a concept of (non-)similarity that does not apply to a single individual (or species) but, instead, it requires at least two entities. This generalization implicitly defines the level of detail and dimensions of comparison. The level of detail may even change along a sliding scale depending on the evolutionary affinity of the individuals being compared: comparing individuals within a species would require consideration of different variables than comparing different species within a genus.

Species area curves summarized as $S = CA^{z}$ (where S is species richness, C is a constant, A is area sampled, and z is scaling constant) or allometric relationships such as $D_i \propto M_i^{-34}$ (where D_i is density of species *i* and M_i is its mass) are good examples of other empirical generalizations that commonly function in ecology (Gould 1979, Whittaker 1998), in spite of doubts as to their validity (e.g., Scheiner 2004). One characteristic indicator of generalization is the (initial) absence of theory that is capable of reproducing the underlying pattern without exceptions.

Usually, an empirical generalization provides a strong stimulus for finding a conceptual framework that is capable of providing an explanation for the empirical pattern detected. This was certainly the case with S=CA^z for which Preston (1962) developed statistical explanation and the equilibrium theory of island biogeography provided a additional biological mechanisms. Also, currently, the metabolic theory of ecology (Brown et al. 2004) proposes answers for the observation that $D_i \propto M_i^{-34}$.

Idealizations on the other hand are theoretical devices created to explore consequences of a feature or a relationship of interest as if no other interfering factors were at play. To continue with an earlier example, one could assume that species are randomly distributed in homogeneously diverse space (space in which a mix of features repeats itself over and over; cf. Hutchinson 1961). Then one could derive a formula describing the relationship between

diversity and the size of the area sampled as Preston (1962) did. If one further assumes that species abundance distribution is log-normal, this idealized relationship would produce $S = CA^z$, independent of historical events, uneven habitat availability, and the presence of dispersal barriers. In the process of idealization, habitat properties and species traits are rendered irrelevant. Such idealizations are commonly used in quantitative (unstructured) models where, for example, a population may be viewed as a collection of identical particles all with the same set of attributes (same probability of reproduction, death, resource acquisition, and others, selected as convenient). Similarly, ecosystem models employ a common currency (carbon, calories) to represent movement of materials and energy among components. Each flow, however, being different from other flows in the real world, is thus stripped of all properties save its currency value to create an idealized representation of the ecosystem.

A BRIEF TRIP INTO THE PAST

Ecology made a much greater use of notions, and generalizations than of formal models and broader theories. Nevertheless, all were combined over time into a loose mix that insured a considerable growth of understanding and conceptual sophistication. At the beginning, ecology was strongly influenced by other sciences, from which it grew, borrowed, or was inspired by (McIntosh 1985).

Although the Malthusian geometric population growth model was not developed by an ecologist, its implications influenced Darwin's ideas on the selective pressure of the environment. Darwin brought an ecological process to bear on evolutionary considerations. He may not have cared about the distinction between ecology and evolution, nor even exposed to Haeckel's later coinage of the term "ecology," nonetheless he recognized that the ecological stage is essential for the evolutionary play to proceed. A mathematical model had undoubtedly contributed to the emergence of the most powerful theory of biology, evolution by natural selection, with continuous and deep implications for ecology.

A little later, Forbes (1887) published his oft cited paper, "Lake as a Microcosm". Although he may not have been the first to do so, his paper makes use of a number of important concepts that soon became, and still are, at the core of the ecological paradigms. First, and most importantly, Forbes saw that interacting organisms in a habitat form a system. When he referred to predator-prey interactions as an example, he used the phrase "a close community of interests

among these ... deadly foes". Second, by emphasizing the relative autonomy of lake biotic and abiotic processes, he identified another important aspect of ecology - patchiness or partial discreteness of ecological phenomena. These ideas were later pursued in ecosystem (Golley 1993), patch dynamics (Pickett and White 1985), and succession studies (Clements 1916), among others. I note here that Scheiner and Willig's framework (Chapter 1) does not explicitly address the discreteness of ecological things and processes. Perhaps this is not necessary or perhaps the proposed framework will find room for such aspects of nature later on.

Clements' ideas of a community of species acting as a superorganism are now deemed a failure (Hagen 2003). However, Clements' ideas had a considerable pre-theoretical and theoretical content, which provoked and inspired ample empirical work that, ironically, largely intended to disprove his superorganism perspective. Ultimately, a more individualistically oriented interpretation of species assemblages won (now also known as the Gleasonian view). Nevertheless, Clements infused into ecology notions that continue to raise their head for good reasons. Although ecological systems do not behave like organisms, they have features that they share with organisms, even if the expression of these features is much less prominent. Forbes preceded Clements in emphasizing interdependence of components, boundedness, and the equilibrial nature of the ecological systems. Clements' and Forbes' perspectives on communities were pre-theoretical. However, it is Clements who made as strong a case as it was then possible, and who left ecologists thinking about these issues for a good while. Most recently, Loreau et al. (2003) returned to some of the issues (e.g., component interdependence, conceptualization of ecosystems as interacting entities) in a modern way, thus completing another cycle of refinement of theoretical thinking. Evolutionary ecologists (e.g., Wilson 1997) emphasize that coevolved species have the potential to form communities with meaningful integration (and thus a degree of entitization). Thus, while the ideas of Clements on the organismal nature of communities may not be applicable to a single trophic level such as assemblages of plants, they are far from irrelevant as noted by Tansley (1935).

However the argument about integration may unfold, it is rather clear that a meaningful consideration of many ecological processes requires system identification. Often such identification remains as an implicit assumption as it has been, for example, during much of the development of theory of succession. Only recently (Pickett et al. Chapter 9; Pickett et al. 2009)

a systematic analysis of the structure of that theory led the authors to augment the model of succession with the concept of system boundary. Bravo!

These observations would be incomplete without commenting on the role of competition and predation models, proposed independently by Lotka (1925) and Volterra (1926). From today's perspective, these models represent more than clever formalizations. Indeed, they have been successful in advancing a field of primarily two-species interactions, in addition to inspiring food web modeling (e.g., Cohen 1978) and the examination of the effects of diversity on stability (e.g., Pimm 1980). More importantly, they led ecology by showing that complex processes can be distilled into manageable characteristics whose behavior can be examined using mathematical tools. In doing so, these models inspired many other quantitative forays into ecology. It may not be possible to make direct connections between these models, but their impact on the way ecologists think is undeniable. Every general textbook ecology presents and explains these models, their form lends itself to analytical solutions and numerical simulations, their generality permitted adding realistic terms (self-limitations, time delays, stochasticity, patchiness) to explore specific factors, and overall, they contributed greatly to the development of quantitative ecology (Silvert 1995).

Landmark theories and concepts appeared in ecology only sporadically. The major ones are shown in Figure 2.2 and Table 2.1. Although I recognize that Figure 2.2 is incomplete in its coverage, and that different ecologists would be likely compile different lists of theories and concepts, two tentative observations come to mind. First, there was a considerable acceleration in the number of major ideas, often well formalized, between 1930 and 1970. Second is the slowdown in the appearance of new propositions. I do not interpret this slowdown as a real problem. The period in question witnessed tremendous development of ecological theory and its penetration into all facets of ecology, from the language that ecologists use to the design of field research. This growth occurred within the broader theories established earlier, through their refinement, expansion, testing, or application in conservation or management. Chapters in this book provide numerous examples of the continuing progress along those lines. Thus, Figure 2.2 is by no means to be interpreted as evidence of limited growth of ecological theory. The growth has been impressive. However, this growth may have unintended consequences, such as the fragmentation of ecology into subdisciplines dominated by their own concepts, theoretical

constructs, and conventions. Finally, it is possible that the newest theories not listed here will need time to register widely in the discipline of ecology before they can appear in Figure 2.2.

The history of theoretical developments in ecology involves more than the timing and appearance of those constructs, or their direct effects on the discipline. In the long run, these constructs effectively define ecology, its domain, questions, and directions. They do so in a rather unsystematic manner. Nevertheless, the nature of ecology would be very different if not for the various relationships that theory has with the rest of ecology.

THEORY SUPPORTS ECOLOGY

I believe that ecology has been molded by theory to a much greater extent than commonly recognized. Although this thesis may appear trivial to those interested in theoretical progress and the unification of ecology, a great number of ecologists express some skepticism and even scorn of theoretical endeavors. Revisiting older and more recent developments and arguments shows that theory plays a central role in stimulating ecological research. Indeed, functions and relationships of theory in science are numerous. Some of these functions and relationships can be identified in ecology. Below, I comment on several common facets of the use of theory in ecology, from historical patterns to current evidence for the contribution of or need for theory.

These comments focus on ways in which theory stimulates, guides, or assists empirical research in some other ways. Under separate subheadings, I provide examples of situations where theory (e.g., mathematical interpretation of species abundance patterns) inspired efforts to accumulate new observational cases in order to verify patterns suggested by that theory, where broad empirical patterns were generalized (allometric relationships) and tentatively explained by theory (metabolic theory), where a new theory (neutral theory of diversity) led to many tests of its assumptions and predictions, or where early ideas (succession) underwent several cycles of refinements and empirical challenge. I also comment on the practical benefits that theory provides, whether by helping with the design of experiments or by providing an intellectual reassurance for conducting research. Because not all theories that ecology uses have been initially formulated in ecology, I note some links that ecological theories have with other disciplines, and the theory of evolution in particular.

Theory, although an indispensable vehicle for generating understanding and organizing knowledge, occasionally hindered progress, at least in other sciences. I note at least one

circumstance related to theory that may result in an unnecessary hurdle on the path of advancement of ecology: excessive emphasis on mathematical models at the expense of effort to formulate a general and unifying theory.

Empirical efforts arise in response to theory

Historically, much empirical research was a response to natural history questions and to theoretical propositions, even if the latter represented unsuccessful attempts to capture and explain observed patterns. Species abundance curves provided a richness of examples. The observation was that most collections of species representative of a local community are characterized by a particular abundance distribution: a few species are abundant, many are rare, and some are in the middle of the abundance range. Initially, this observation was reported in the form of an empirical generalization summarized in the form of a mathematical distribution. Preston (1948) proposed that the abundance of species in a community, when plotted on the log scale, shows a hump and generally fits a log-normal distribution. Later on, the focus shifted to mathematical models with biological underpinnings. MacArthur (1957), Whittaker (1965), May (1975), Tokeshi (1993), and many others proposed a bewildering selection of models mostly based on assumptions about partitioning of resources (e.g., broken stick model), with resources being logically tied to abundance. Others invoked habitat structure (Kolasa 1989) or population dynamics (He 2005) to explain species abundance distributions. Ecologists have been and continue collecting data to test (fit) those models.

Broad empirical generalizations beg for theory

The role of empirical generalizations is different than that of theoretical frameworks or models. Perhaps one of the more important aspects of broad empirical generalizations is that they provoke questions about causes of patters and, as a result, give birth to theoretical explanations and subsequent tests. Examples are abundant in most ecology textbooks. Consider allometric relationships and their impact on theory and, more importantly, the continuous interests in those relationships because of theory. Scientists have known for nearly two centuries that larger animals have slower metabolisms than smaller ones. A mouse must eat about half its body weight every day to avoid starvation; a human gets by on only 2% (Whitfield 2004). This coarse empirical generalization, like others, begged for explanation. The first theories to explain this

trend were based on the ratio between body volume and surface area. The square-versus-cube relationship makes the area of a solid proportional to the two-third power of its mass, so metabolic rate should also be proportional to mass. But a thorough study by Kleiber (1932) found that, for mammals and birds, metabolic rate was mass to the power of 0.73 (approximately three-quarters) and not 0.67 (two-thirds as postulated based on surface to volume ratios). Other research supported this new three-quarter-power law, although consensus is still elusive (Grant 2007). It was, however, much harder to find a theoretical reason why metabolic rate should be proportional to mass raised to power of 3/4. Furthermore, it was not clear why quarter-power scaling laws should be so prevalent in biology. Nevertheless, the generalization and the difficulties with the putative causes of it led West et al. (1997) to develop a new explanation of why metabolic rate should equal the three-quarter power of body mass.

Indeed, the generalization that larger organisms have lower metabolism is such a noisy one that some question its relevance. What is important is that an early strong empirical generalization inspired development of sophisticated theoretical explanations. A considerable value of these explanations to ecology arises not from whether they are correct or accepted but rather whether they have been formulated and tested. The very process of doing so identifies further problems, questions, and directs new empirical and theoretical pursuits; which advances understanding.

Another theory formulated in response to an empirical generalization is the neutral theory of species diversity (Hubbell 1997). Most communities, irrespective of the scale of data collection, harshness of the environment, or productivity, show a general qualitative regularity of species abundances: few species are abundant while many more are rare, with species of intermediate abundance filling the range in between according to one of several abundance partitioning models. Numerous theoretical efforts have been made in the past to account for this general pattern (e.g., May 1975, Ugland and Gray 1982, Tokeshi 1993). The most familiar of those is the broken stick model of MacArthur's (1957). None have gained full acceptance. Hubbell's proposition is a culmination of these efforts and a stimulus for further tests. Admittedly, sometimes a good generalization does not lead to the development of a theory (but see Fox et al., Chapter 14, on gradients regarding a possible solution). Holdridge's schematic, illustrating vegetation types as a function of evapotranspiration and annual precipitation (Figure 2.3), represents a useful generalization that does not need or inspire a complex theory. In case of

vegetation types, a simple combination of non-biological variables and a few physiological assumptions together provide an excellent explanation of the patterns. This example illustrates further that almost all efforts to systematize ecological information, including generalizations, involve some theoretical content, a point that I already presented earlier.

35

New propositions point toward new fruitful directions

A fairly recent 'metabolic' theory (West et al. 1997) of ecology witnessed an intensive activity (over 700 citations) that concentrates on two major foci: the establishment of empirical generalizations and the testing of assumptions. While attracting a fair share of praise, it has proved a magnet for criticism as well (Grant 2007). Perhaps one of the most stimulating features of this theory has been its direct predictive power. Predictions are relatively easy to derive and test. The mathematical relationships at its core seem to make it particularly amenable to and the immediate reason for the testing. Few propositions with a broad scope and potential generality have this advantage. However, it is too early to make conclusions about the acceptability and ultimate place of this theory in ecology as critics point out to various difficulties (e.g. Li et al. 2004 and other comments in Ecology's Forum; Kozłowski and Weiner 1997), but its influence on current research remains high as demonstrated by the number of papers addressing various aspects of the theory, from empirical patterns to nuances of assumptions.

Old propositions continue to mature and be refined

Succession theory is a prime example of the importance of theory to ecological research and its progress. It started with the simple observation that vegetation changes over time. First attempts to record and explain the patterns of change are due to Cowles (1899) and Clements (1916). These attempts led to more empirical research aimed at verifying the patterns or contesting early generalizations (e.g., succession leads to a climax; Connell 1972). The debates and new data culminated in a sequence of models of succession, of which Connell and Slatyer's (1977) tolerance-inhibition-facilitation model has become common textbook knowledge. Once these more logically precise models were available, it was possible to expose their weakness and to attempt "a professional" grade theory formulation (see Pickett et al., Chapter 9, for an in-depth account).

Experiments invoke theory

In addition to the impact of well defined, broad scope propositions such as succession theory, island biogeography theory, or metabolic theory of ecology already mentioned, theory affects ecology in many indirect ways. As ecology expanded its use of experiments, theory was one of the most common motivations for experiments. While attributing credit to theory may be difficult for empirical generalizations, particularly those accumulated from numerous cases over long periods of time, experiments almost invariably find roots and rationale in theory. Whether this is a general theory, a specific model, or even a vague model, depends on the preferences or awareness of the authors of the experiment. It might be difficult to find a contemporary paper that presents experiments that makes no reference to some theory.

Theories provide comfort

Most ecologists are familiar with, or at least heard of, population growth models, foraging theory, competition models, metapopulations, niche theory, succession theory, or island biogeography theory. This is because much of ecological research is either motivated by these theories and models, or their specialized offshoots, or that they are used to interpret the results of empirical studies. Theory often provides a compass to studies whose primary purpose may be gathering or compilation of data. In this manner, theory also provides reassurance as to the validity and significance of efforts by offering a context within which to make strategic research decisions. Leonardo da Vinci was right. Without these and some newer or more specialized theories, ecological research would not make much sense beyond some practical cataloging of observations.

General science theories enrich ecological frameworks

Examples of general science theories that enrich ecological frameworks include systems theory and biogeography. In 1982 Allen and Starr (1982) published a book entitled "Hierarchy". It was meant largely as a methodological book in the broadest sense. The book refined and recast general scientific understanding of how the world is constructed and perceived by scientists, with a special focus on ecology. Although the book presented no quantitative formulations or theory per se, it emphasized the inevitability of changing interpretations of the observed data sets as a function of scale. Its impact (enhanced by a couple of companion books with Allen as co-
author) was impressive. Although few studies tested directly premises that form the constellation of ideas jointly called 'hierarchy theory', Allen's writings and conference presentations made ecologists aware of the effects different observational scales may have on the interpretation of results of their studies. Today (2008) more than 500 papers show up in searches with keywords "ecology" and "scale" and provide evidence for the extensive impact made by this theory (Schneider 2001). In this book, Peters et al. (Chapter 13) draw substantially on the ideas presented by Allen and Starr not so long ago. MacArthur and Wilson's (1967) theory of island biogeography had a similarly stimulating impact. Not only did it encouraged numerous tests which, incidentally, illuminated the importance of a number of additional factors such as island heterogeneity, differential colonization leading to compositional disharmony, or incomplete saturation, but also became a major conceptual ingredient of conservation strategies (Sax and Gaines Chapter 10).

Evolutionary theory led the way

The theory of evolution has been tapped by ecologists from the beginning, and has provided a sound and rich framework for the studies of habitat selection, parental investment, trade-offs, behavioral ecology, just to give a sample of the range of problems addressed. However, the main and readily identifiable signs of the link between the theory of evolution and ecology appeared relatively late (e.g., Lack 1954, Brown and Wilson 1956, Hutchinson 1965). The existence, caliber, and continuous impact of important journals largely devoted to the exploration of the interface between evolution and ecology (American Naturalist, Evolutionary Ecology Research) testifies to the vitality of evolutionary theory in ecology. It further adds to the weight of the argument that theory plays a fundamental role in advancing ecology, even when ecology cannot claim the primary ownership of that theory.

Theory is a vehicle for sharing knowledge across ecological sub-disciplines

I offer an informal observation that ecologists share important information across sub-disciplines by using theory. Although this sharing may involve simplified versions of theory, it is nevertheless the main material that glues the science of ecology together. For example, whether one conducts research in plant physiology, ant behavior, nitrogen pathways in soil, reproduction of deer, squirrel allocation of time to foraging versus caching, clines of diversity, or global

carbon cycle, one is likely to be familiar with most theories and concepts listed in Figure 2.2. Furthermore, theory is often seen as an agent for dissemination of ideas among subdisciplines. Ecologists working on streams, soils, or other habitats reach to general theories for inspiration and guidance to questions and problems and adopt the concepts and definitions associate with such theories. For example, Lake et al. (2007) explicitly call for linking ecological theory with stream restoration. Similarly, Barot et al. (2007) identify the need for soil ecologists to make greater use of evolutionary theory and modeling in order to shift emphasis of soil ecology from particularities of empirical observations to generality that they associate with ecological theories.

Ecologists often think of theory as mathematical models

A strong debate developed in the 1950s and 1960s among the proponents of theory in ecology based on the response of organisms to resources and habitat conditions (Andrewartha and Birch 1954) and those who saw the need for inclusion of evolutionary and community processes (Lack 1954, Orians 1962). Battles for the conceptual vision of ecology, for its domain, main assumptions about the subject matter, and the structure of theory are important. Many research programs were undoubtedly influenced by arguments arising in the course of such debates, whether a specific debate was about the inclusion of evolutionary processes, energetic, or non-equilibrium perspective. However, books devoted to theoretical ecology (e.g., Case 2000, Roughgarden 1998, Yodzis 1989) make no mention of this or similar conceptual debates. These texts present theory as collections of mathematical models. Although such models are powerful and illuminating theoretical constructs, their very dominance of the theoretical landscape of ecology may have unintended consequences because it may detract from or undervalue the significance of the efforts to reorganize the conceptual framework of ecology. Fortunately, the project initiated by Scheiner and Willig may mitigate this potential negative effect.

Indeed, there were earlier attempts to refine this framework. Schoener (1985) attempted to accommodate the diversity of models and perspectives by calling for 'pluralistic ecology'. Restricting his scope to community ecology, Schoener believed that the best approach was to develop separate mathematical models using six primitive (e.g., body size, motility), six environmental (e.g., spatial fragmentation, severity of physical factors) and six derived (e.g., relative importance of competition and predation) axes to classify collections of species into different community types. Different types of communities would then be approached as

separate theoretical problems that require separate treatment that lead to separate solutions. Others (Reiners and Lockwood 2009) extend this view to all of ecology. I am not convinced however that the best course of action in the face of conflicting theoretical positions is a compromise of this kind (see also McIntosh 1987). A plurality of approaches may represent a necessary stage but little precedent supports the idea that will offer an effective solution to conceptual mix that ecology offers today. Also, it is likely that pluralism or fragmentation of the domain and methodology will impede discovery of and consensus about empirical generalizations.

EMPIRICAL DATA SPURS THEORY

In contrast to the previous section, empirical research may also be motivated by theory or condensed to form generalizations. Generalizations call for explanations, which are provided by theory, and so the cycle continues. However, empirical discoveries may sometimes open new avenues of inquiry independently of the influence of any particular theory. The discovery of hydrothermal vents and their associated communities (Lonsdale 1977) had significant implications on ideas about the origin of life (Wächtershäuser 1990), on perspectives concerning food chains, and affected thinking in the new field of exobiology. Records of pest outbreaks pose a continuous challenge for theoreticians and stimulate both further empirical work and theoretical approaches (Stone 2004, Dwyer 2004). Examples where new observations and discoveries spurred a flurry of research activity are numerous (interspecific carbon exchanges among plants, coral bleaching, acid rain, morphological changes in zooplankton in the presence of predators, appearance of zebra mussels in North America, and many others). However, the meaning and value of such discoveries were quickly enhanced when wedded to a theoretical construct (generalization, working explanatory hypothesis, or broader theory). On their own, empirical findings, unlike theory, rarely give impetus to a robust program and steady research direction.

DIFFICULTIES

In spite of numerous examples of theory being important to ecology, the direct use of theory is still modest and, perhaps, limited by tradition, training, or skepticism. In 2008, the November and December issues of Ecology contained a total of 57 articles and comments of which only 15

were clearly motivated by theory or formal models. Many papers did not invoke theory to any extent. However, even the most empirically oriented papers were firmly seated within modern paradigms of ecology, which itself is strongly shaped by theoretical views and constructs. The American Naturalist favors theoretically justified papers to a much greater extent, although many papers in that category involve evolutionary models. For example, the November and December issues in 2008 contained 32 papers with 25 as primarily motivated or driven by theory. The 1929 (two issues each) of these two journals had almost no papers involving ecological or evolutionary theory; only 3 of 28 in Ecology and 1 of 21 in American Naturalist were expressly motivated by theory. A search of JStor (Ecology and Environment section) shows that the number of papers with keywords (see legend to Figure 2.4) that indicate theoretical motivation or links tends to increase at a faster rate than the number of papers without such a designation (Figure 2.4; cf., slope values). Whatever the particular trends, the influence of theory on ecology grows quickly and, very likely, faster than that of the observational component.

The growth of influence associated with the depth of analyses, improved expertise in translating theory into lab and field research, as well as the rise of many specific theories continues to contribute to the splitting and fragmentation ecology into areas with rather independent existences. Growth is good, but fragmentation introduces and nurtures conceptual incongruities that hamper ecology in my view. In support of this idea, I cite two examples: one explicitly identifies the deficiencies due to conceptual isolation while the other shows how unification and progress could be achieved if cross-fertilization of ideas took place. For example, a considerable debate developed about the nature of competition among plants. Specifically, the relationship between nutrient availability as affected by potential competitors, allocation to root growth, and consequences of these factors for plant growth and reproduction developed as major concerns among plant ecologists (Craine 2007). Yet, specific answers are unlikely to interest animal ecologists and the conceptual refinements associated with them are unlikely to apply beyond plant ecology. The second example concerns two separate research traditions of parasite-host and predator-prey ecology. Raffel et al. (2008) argued that developments in predator-prey ecology, such as temporal risk allocation and associational resistance, can contribute to development of new hypotheses for parasite-host systems. Conversely, concepts developed in parasite-host ecology, such as threshold host densities and phylodynamics, might enrich for predator-prey ecology. Propositions such as trait-mediated

indirect effects and enemy-mediated facilitation provide opportunities for the two fields to forged a shared theoretical perspective that would foster advances in both fields (Raffel et al. 2008).

This troubling and somewhat perennial state of affairs motivated several attempts at unification. At various times, theories that we now recognized as of limited domain were advanced to fulfill that role. In the 1960s, ecosystem theory attempted to explain all by ignoring detail and by focusing on physical and chemical fundamentals and constraints. At the same time, population ecologists thought of extending the mathematical framework of population ecology to community ecology, macroevolution, and biogeography (Odenbaugh Chapter 3). More recent examples are the neutral theory of diversity and the metabolic theory of ecology discussed above. There were others with a lesser bang but still offering evidence of interest (e.g., Belgrano and Brown 2002) in bringing more cohesion to the fragmented science. Clearly, ecology itches for a general theory. Unfortunately, so far ecology lacks a well articulated general theory and even a general framework for relating available theories (but see Scheiner and Willig, Chapter 1, that aims to lay a foundation for the erection of such a general theory by a thorough examination of the domain and foundational principles).

Perhaps this last obstacle, i.e. the lack of a general theory, will come as no surprise. Ecology is focused on the natural world and as a science that draws on the passion of those who love to study it. As such, it does a limited job at preparing young scientists in the use of theory. This tradition results in ecologists being more familiar with the requirement to formulate testable hypotheses but less familiar with identifying the domain, assumptions, central relationships, or the developmental status of various components. Consequently, ecology suffers from a level of resistance to new theoretical propositions (entrenched paradigms?), particularly those derived from traditions other than its own (e.g., the skeptical reception that hierarchy theory received initially; e.g., Keitt 1999). Fortunately, the eloquence and persistence of its main proponents (T.F.H. Allen and R. O'Neill) led to a wider acceptance of some of the theory's premises (e.g., pattern shifts with scale) and benefited research in subsequent years, if the benefit is measured by the number of times the term "scale" appears in the literature (Schneider 2001). The resistance to theory may come also from apparent success of descriptive or quasi-descriptive work. As many studies uncover and describe important facts of nature by simply asking "why" or "how", they ease the urgency of using theory and of building a synthetic theory of ecology. The attitude that "It's all very well in practice, but it will never work in theory" (a French proverb) may be

instantaneously satisfying but a serious hindrance in the long run.

CONCLUSIONS: PROS AND CONS OF THIS REALITY

Pros - clarity of focus and economy of effort

The benefits of organizing research within a theoretical framework or by directing research towards a prospective framework are fairly well understood (see Pickett et al. 2007). Theory assists researchers in identifying appropriate questions, provides guidance as to what approach will be suitable to test the hypotheses, and offers criteria to judge the progress in answering these questions. At each of these tasks, theory helps in making specific decisions or assessments.

By streamlining and directing research, theory is the only guarantee, although not infallible, that the journey from observation to question and to the best available answer is as short as possible. Of course, wandering minds may by chance find shortcuts and do better than will a researcher who systematically uses theory. However, wandering minds offer no guarantee of success and, more often than not, travel in unproductive directions.

Cons – exclusion and barriers to ideas not sanctioned by current theory

Any established framework resists change. Although the current body of ecological theory is an aggregate of loosely related and unrelated constructs, it still offers substantial resistance. This resistance may come in the form of marginalization of some research avenues by disproportionate emphasis on others (e.g., the International Biological Program has been dominated by trophic/energetic concerns; Golley 1993), misinterpretation of the promise new propositions may offer, or falling back on a minimalistic agenda. For example, a search for general laws of ecology may be discouraged because of skepticism as to whether they can be found (Lawton 1999, Knapp et al. 2004).

It is possible that in being very successful, Lotka and Volterra delayed other theoretical developments by shifting much of ecologists' attention to the application of differential equations at the expense of other approaches to formalizing ecological processes. I am not arguing that this happened, but rather that our praise of their contribution may be unbalanced, and that the matter needs careful consideration in the future.

Misunderstandings may stimulate empirical research, which ultimately may lead to an alignment of theory and observations. When Robert May (1973) published his analysis of the

effects of diversity on the stability of ecological systems, many were puzzled, but appeared to accept the result that diversity is destabilizing (McCann 2000). Specifically, by using linear stability analysis on models constructed from a statistical universe (i.e., randomly constructed communities with randomly assigned interaction strengths), May found that diversity generally destabilized community dynamics. However, May's results did not imply at all that diverse communities should be unstable, rather he showed that communities comprising random species should be. This crucial difference was overlooked for a couple of decades. Furthermore, May included no modularity in connecting species and treated the interactive network as a single unit. More recent research indicates that this makes a major difference because modularity insulates unstable species groups from each other (e.g., Kolasa 2005, McCann 2005). Nonetheless, May's approach and results stayed at the center of discussions on the role of diversity in maintenance of ecosystems and their properties. Importantly, May's articulation of the problem led to further theoretical and empirical work that continues to provide valuable insights regarding the role of species diversity in ecological systems. In this case, misunderstanding put ecology on a wrong path from which it recovered gracefully and with the benefit of stimulating an important area of inquiry.

Future: need to integrate and use theory more thoroughly

Charting the course for the future may be made easier by looking at types of theoretical work and their relationships to the different goals ecologists choose to pursue, whether it is generality, precision, or realism (Figure 2.5). Although these goals are not mutually exclusive in principle, in practice, theoretical constructs show more strength in addressing one of the goals than others. Without pretending any analytical vigor, I would like to register some impressions. Theoretical constructs in ecology (models, low level theories, or even generalizations) fall somewhere between realism and generality (Figure 2.5). I suggest this is so as a compromise between their features: many models such as those based on population dynamics (predator-prey, competition, metacommunity) are general in formulation. This generality contrasts, however, with their limited scopes. Take a logistic population growth model. The model can be applied to any population – it is thus general. However, a population is only a tiny fraction of the ecological universe. The model does not necessarily apply to the whole species as its carrying capacity and genetic constitution may vary in space and time. Moreover the model does not apply to family

groups, colonies, or symbiotic systems, or to coevolved multispecies systems. From this perspective, the model is not general at all. It is just one in a collection of complementary, and sometimes interacting or overlapping, models. By having a narrow scope, the model can be reasonably realistic if parameters are estimated for a limited natural universe. Still, models of populations, or interactions among populations in different contexts, are a major stronghold of theory in ecology.

By contrast, ecosystem models are often developed for specific natural systems (Fitz et al. 1996). Such models rely on general physical principles to account for the energy and material budgets and on biological relationships to provide system specific and local contents. They can be quite realistic, at least this is the intention of the investigators, but attempts to make them general would require stripping biological contents and so removing these models from the domain of ecology. In any case, I believe that ecology has benefited from developing and accumulating ecosystem models in several ways such as building bridges to cybernetics and engineering, providing management tools (e.g., adaptive management – Holling 1978), and explicitly incorporating the abiotic environment into the dynamics of ecological systems.

Still another category of models appears to be driven by the need for precision. In an attempt to explain species rank-abundance distributions ecologists developed a number of statistical models. They are judged by the precision with which they describe actual collections of data. In the absence of a good theory about underlying mechanisms, these models cannot be seen as general. In fact, these models seem to work well in particular situations only, for example, large data sets from undisturbed habitats generally follow a log-normal model of abundance distribution. Different may models apply to disturbed habitats or small data sets. As these models do not make good explicit links to underlying processes, they can hardly be seen as realistic, even when they provide a precise fit to the data.

It is surprising, and it holds some promise, that some of the early, potentially important, concepts have not advanced theoretically for over 100 years. Forbes' microcosm embodied the idea of wholeness and relative autonomy. He articulated, early on, a central notion that the ecological world exists in the form of aggregates of interdependent parts. In his view, this aggregation would be responsible for the emergence of partially bounded systems whose components with a appreciable ability to adjust internally to each other. It is rather surprising that this idea, while often invoked in the context of conservation, has not developed theoretically

to any major extent. Although ecosystems ecology attempted to deal with properties of entitylike aggregations of organisms, it now appears to be stuck in narrow mechanistic representations of flows and budgets (e.g., Potter et al. 2001, compare boreal ecosystem models with respect to annual carbon and water fluxes to evaluate carbon budget). One of the biggest and most interesting ideas of early ecologists is still a promising area for theoretical development.

45

The recent proposition by Scheiner and Willig (Chapter 1) that is at the heart of this book will undoubtedly stimulate and provoke further work. It will mature, evolve, lose unnecessary components, or relate them more efficiently. From the perspective of a practicing ecologist, the eight principles they propose are difficult to employ because the relationships among them are not yet succinctly established. Also, an important element may be still missing. I address this and identify the element in the next section.

Future: we need to define the domain of ecology and the level(s) of organization

Ecologists have traditionally been rather reluctant (or oblivious) to the problem of domain. However, clarity as to the domain of a proposition is a sine qua non condition for evaluating a proposition. Scheiner and Willig do a good service (Chapter 1) to ecology in the way they tackle the question of domain – they comment on and assess domains of various constituent theories of ecology and their overlaps or exclusions with other sciences. In the process they show that currently available theories cover portions of the entire domain of ecology. This is a work in progress, but its completion may help to consolidate the theory of ecology.

A unified theory of ecology must, satisfactorily, deal with several levels of organization. Some believe that a successful theory of ecology must be based on the distribution of organisms (including its causes and consequences, e.g., Gleason, Andrewartha; Andrewartha and Birch 1954; Wiens 1984; Scheiner and Willing, Chapter 1). Others take an even more radical position and claim that natural selection at an individual level is a sufficient explanation, and thus a basis for a general theory, of all ecological phenomena (Williams 1966). Still others, with Clements, Odum, Wynne-Edwards, Patten, Ulanowicz (1997), or Jørgensen (2007) emphasize entities arising from interdependence of species (and environment) such as communities and ecosystems to be the central objects of interest to a general theory. All of these ecologists make good arguments but the troubling problem is that these arguments appear incompatible at the moment.

Two general strategies may hold promise in solving the century old split. One is to

gather existing theories in a unifying superstructure. This is the approach taken in this book. While it does not specifically address the problem of multi-individual or multi-species entities, in principle the hierarchical structure of a unified general theory has the capability of accommodating many well articulated lower level theories. Alternatively, a unified general theory that does not center on one or on narrow range of organizational scales may be capable of resolving the conflict as well. But such a theory may need another book.

In any case, it may be useful to add one more principle (to add to eight identified already by Scheiner and Willing in Chapter 1) that reflects a generalization that all theories of ecology, and thus the prospective general theory of ecology, deal with ecological entities. Ecological entities appear at different levels of organization: they may be individuals, kin groups, local populations, metapopulations, symbiotic systems, local communities or ecosystems (e.g., Wilson 1988), and many others. Scheiner and Willig are clearly aware of this when they mention variation among individuals – a label that includes individual organisms and individual species. However, different degrees of entitization constitute a universal feature of living nature. An explicit recognition of it as a ninth principle might aid in advancing the general theory of ecology.

IN A NUTSHELL

- Theory has had strong direct and indirect influence on the paradigm (mindset, culture, standards, and directions) of ecology.
- Ecology without theory would be a science accumulating cases, without the ability to develop and evolve, and without the ability to make sense out of the multitude of cases, except in the light of theory of evolution.
- Even though, the systematic pursuit of theory has been modest, the successes have largely been idiosyncratic due to individuals who addressed particular questions without an overarching framework.
- Awareness of theory and its use appears to increase with time across the discipline, but weaknesses persist.
- In spite of efforts to remedy the situation, ecological theory has until recently comprised theories and models of narrow scope.

• A larger, integrating theory is lacking, but promising efforts towards its formulation are underway.

Acknowledgements

I thank Sam Scheiner and Mike Willig for creating an opportunity to reflect on historical and contemporary issues of general ecological theory. Two anonymous reviewers generously helped to clarify my ideas and their presentation.

Literature Cited

- Allen, T. F. H. and T. B. Starr. 1982. *Hierarchy: perspectives for ecological complexity*, Chicago, University of Chicago Press, 310 pages.
- Andrewartha, H.G., and L.C. Birch. 1954. *The distribution and abundance of animals*. Chicago: University of Chicago Press.
- Barot, S., Blouin, M., Fontaine, S., Jouquet, P., Lata J-C, and J. Mathieu. 2007. A tale of four stories: soil ecology, theory, evolution and the publication system. PLoS ONE 2(11): e1248.
- Belgrano, A. and J.H. Brown. 2002. Ecology: Oceans under the macroscope. Nature 419:128-129.
- Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., and G.B. West. 2004. Toward a metabolic theory of ecology. Ecology 85: 1771-1789.
- Brown, W. L., and E. O. Wilson. 1956. Character displacement. Systematic Zoology 5:49-65.
- Camerano L. 1880. Dell'equilibrio dei viventi merce la reciproca distruzione. Atti della Reale Accademia delle Scienze di Torino, 15: 393-414.
- Case, T.J. 2000. An illustrated guide to theoretical ecology. Oxford University Press, 449 pages.
- Clements, F.E. 1916. Plant succession: An analysis of the development of vegetation. *Carnegie Institution of Washington*, publication no. 242.
- Cohen, J. E. 1978. Food webs and niche space. Princeton, Princeton University Press, 189 pages.
- Connell, J.H. 1972. Community interactions on marine rocky intertidal shores. Annual Reviews of Ecology and Systematics, 3:169-192.
- Connell, J.H. and R.O. Slatyer. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. American Naturalist 111:1119-1144.
- Dwyer, G., Dushoff, J. and Yee, S. H. 2004. The combined effects of pathogens and predators on insect outbreaks. Nature, 430:341–345.
- Elton, C. 1927. Animal Ecology. Sidgwick and Jackson, London. Fisher, R.A, Corbet A.S.,
 Williams C.B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12: 42-58.
- Fitz, H.C., E.B. DeBellevue, R. Costanza, R. Boumans, T. Maxwell, L. Wainger, and F.H. Sklar.
 1996. Development of a general ecosystem model for a range of scales and ecosystems.
 Ecological Modelling 88:263-295.

Forbes, S. 1887. The lake as a microcosm. Illinois Natural History Survey Bulletin 15:537-550.

- Gause, G.F. 1934. *The struggle for existence*. Williams and Wilkins, Baltimore. Reprinted in 1971, New York.
- Golley, F.B. 1993. *A history of the ecosystem concept in ecology*. Yale University Press, New Haven. 254 pages.
- Gould, S.J. 1979. An allometric interpretation of species-area curves: The meaning of the coefficient. American Naturalist, 114:335-343.
- Grant, B. 2007. The powers that might be. The Scientist 21 (3): 42.
- Grisebach, A. 1838. Genera et species gentianearum. Cotta, Stuttgart and Tubingen.
- Hagen, J.B. 1989. Research perspectives and the anomalous status of modern ecology. Biology and Philosophy, 4:433-455,
- He, F. 2005. Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. Functional Ecology, 19:187-193.
- Holling, C.S. 1959. The components of predation as revealed by a study of small mammal predation of the European pine sawfly. Canadian Entomologist, 91:293-320
- Holling, C.S. (ed.). 1978. Adaptive environmental assessment and management. Chichester: Wiley.
- Hubbell, S. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. Coral Reefs 16 (suppl.):9-21.
- Hutchinson, E.G. 1961. The paradox of the plankton. American Naturalist 95:137-145.
- Hutchinson, E.G. 1965. *The ecological theater and the evolutionary play*. Yale University Press. 139 pages.
- Jørgensen, S.E. 2006-2007. An integrated ecosystem theory. Annals of the European Academy of Sciences, 19-33.
- Keddy, P.A. 1989. Competition. Chapman and Hall. New York, 202 pages.

Keitt, T.H. 1999. Ecological scale: theory and applications. Complexity 4:28-29.

- Kleiber, M. 1932. Body size and metabolism. Hilgardia 6:315–332.
- Knapp, A.K., Sraith, M.D., Collins. S.L., Zarabatis. N., Peel, M., Emery, S., Wojdak, J., Horner-Devine M.C., Biggs, H., Kruger, J., and S. J Andelraan. 2004. Generality in ecology: testing North American grassland rules in South African savannas. Frontiers in Ecology and Environment 2:483-491.

- Kolasa, J. 1989. Ecological systems in hierarchical perspective: breaks in the community structure and other consequences. Ecology 70:36-47.
- Kolasa, J. 2005.Complexity, system integration, and susceptibility to change: biodiversity connection. Ecological Complexity 2:431-442.
- Kozłowski, J. and Weiner, J. 1997. Interspecific allometries are by-products of body size optimization. The American Naturalist 149:352-80.
- Lack, D. 1954. *The natural regulation of animal numbers*. Clarendon Press, Oxford, UK, 343 pages.
- Lake, P.S., N. Bond and P. Reich. 2007. Linking ecological theory with stream restoration. Freshwater Biology, 52:597-615.Lawton, J.H. 1999. Are there general laws in ecology? Oikos, 84:177-192.
- Leslie, P.H. 1948. Some further notes on the use of matrices in population mathematics. Biometrika, 35(3-4), 213-245.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America, 15:237–240.
- Li, B-L., V.G. Gorshkov and A.M. Makarieva. 2004. Energy partitioning between different-sized organisms and ecosystem stability. Ecology, 85:1811–1813.
- Lindeman, R.L. 1942. The trophic-dynamic aspect of ecology. Ecology, 23:399-417.
- Lonsdale, P. 1977. Clustering of suspension-feeding macrobenthos near abyssal hydrothermal vents at oceanic spreading centers. Deep-Sea Research, 24:857-863.
- Loreau, M., N. Mouquet, and R. D. Holt. 2003. Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. Ecology Letters 6:673-679.
- Lotka, A. J. (1925). *Elements of physical biology*. Williams and Wilkins, Baltimore. [Reprinted in 1956: *Elements of mathematical biology*. Dover Publications, Inc., New York, New York].
- MacArthur, R.H. 1957. On the relative abundance of bird species. Proceedings of the National Academy of Sciences, 43:293-295.
- MacArthur, R. H. and Pianka, E. R. (1966). On the optimal use of a patchy environment. American Naturalist, 100:276-282.
- MacArthur, R.H. and E.O. Wilson. 1963. An equilibrium theory of insular zoogeography Evolution 17:373-87.

- MacArthur, R.H. and E.O. Wilson. 1967. The theory of island biogeography, Princeton, Princeton University Press, 224 pages.
- May, R.M. 1973. *Stability and complexity in model ecosystems*. Princeton, N.J., Princeton University Press, 235 pages.
- May, R.M. 1975. Patterns of species abundance and diversity. In: Ecology and evolution of communities, (Cody M.L. and J.M. Diamond, eds.), Belknap Press, Cambridge, pp. 81-120.
- McCann, K.S. 2000. The diversity-stability debate. Nature 405:228-233.
- McCann, K. S. 2005. Perspectives on diversity, structure, and stability. In *Biological paradigms lost* (Cuddington, K., and B.E. Beisner, eds.), San Diego, Academic Press, pp. 183-193.
- McIntosh R.P. 1985. *The background of ecology: concept and theory*, Cambridge, Cambridge University Press, 383 pages.
- McIntosh R.P. 1987. Pluralism in ecology. Annual Reviews of Ecology and Systematics, 18:321-341.
- Möbius, K. 1877. The oyster and oyster-culture. Trans, H.J. Rice. In *Fish and Fisheries, Annual Report of the Commission for the Year 1880*. Volume 3. Appendix H. Documents of the Senate of the United States for the Third Session of the Forty-sixth Congress and the Special Session of the Forty-seventh Congress, 1980-1981, pp. 721-724.
- Odum, E.P. 1969. The strategy of ecosystem development. Science 164:262-270.
- Odum, E.P., and H.T. Odum. 1953. *The fundamentals of ecology*. Philadelphia, W.B. Saunders. 384 pages.
- Orians, G.H. 1962. Natural selection and ecological theory. The American Naturalist 46:257–263.
- Pickett, S.T.A. and P.S. White. 1985. *The ecology of natural disturbance and patch dynamics*. Academic Press, San Diego, 472 pages.
- Pickett, S. T. A., Kolasa, J. and C. G. Jones. 2007. *Ecological Understanding*. *The nature of theory and the theory of nature*, San Diego, Academic Press, 233 pages.
- Pimm, S. L. 1980. Food web design and the effects of species deletion. Oikos 35:139-149.
- Porter, W.P. and D.M. Gates. 1969. Thermodynamic equilibria of animals with environment. Ecological Monographs, 39:227-244.

Potter, C. S., S.Wang, N.T. Nikolov, A.D. McGuire, J. Liu, A.W. King, J.S. Kimball, R.F. Grant, Frolking, S.E., J.S. Clein, J.M. Chen and J.S. Amthor. 2001. Comparison of boreal ecosystem model sensitivity to variability in climate and forest site parameters. Journal of Geophysical Research, 106:671–687.

52

- Preston, F.W. 1948. The commonness and rarity of species. Ecology, 29: 254–283.
- Preston, F.W. 1962. The Canonical Distribution of Commonness and Rarity: Part I. Ecology, Vol. 43:185-215.
- Raffel, T.R., L.B. Martin and J.R. Rohr. 2008. Parasites as predators: unifying natural enemy ecology. Trends in Ecology and Evolution, 23:610-618.

Ramalay, F. 1940. The growth of a science. University of Colorado Studies 26:3-14.

- Real, L.A. and J.H. Brown (eds). 1991. Foundations of ecology. University of Chicago Press. Chicago. 905 pages.
- Reiners, W. and J. Lockwood. 2009. Philosophical foundations for the practices of ecology. Cambridge University Press (in press).
- Scheiner, S.M. 2004. A mélange of curves further dialogue about species–area relationships. Global Ecology and Biogeography, 13:479–484.
- Scheiner, S.M., and M.R. Willig. 2005. Developing unified theories in ecology as exemplified with diversity gradients. American Naturalist 166:458-469.
- Scheiner, S.M., and M.R. Willig. 2008. A general theory of ecology. Theoretical Ecology, 1:21–28.
- Schneider, D.C. 2001. The rise of the concept of scale in ecology. BioScience, 51:545-553.
- Schoener, T.W. 1985. Overview: kinds of ecological communities ecology becomes pluralistic. In: Ecological communities, edited by J. M. Diamond and T. Case, New York, Harper and Row, pp. 467-479.
- Silvert, W. 1995. Is the logistic equation a Lotka-Volterra model? Ecological Modelling, 77:95-96.
- Skellam, J.G. 1951. Random dispersal in theoretical populations. Biometrika, 38, 196-218.
- Stone, L. 2004. Population ecology: A three-player solution. Nature, 430:299-300.

- Strong, D. R. 1984. Exorcising the ghost of competition past: Phytophagous insects. In *Ecological communities: Conceptual issues and the evidence* (Strong, D.R. Jr., D. Simberloff, L.G. Abele, and A.B. Thistle, eds.). Princeton University Press, Princeton, pp. 28-41.
- Tansley, A.G. 1935. The use and abuse of vegetational concepts and terms. Forest Ecology and Management. 16: 284-307.
- Tokeshi, M.1993. Species abundance patterns and community structure. Advances in Ecological Research 24:111-186.
- Ugland K.I., and J. S. Gray. 1982. Lognormal distributions and the concept of community equilibrium. Oikos 39:171-178.
- Ulanowicz, R.E. 1997. Ecology, the ascendant perspective. Columbia University Press, New York. 201 pages.
- Verhulst, P. F. 1838. Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathématique et physique 10:113–121.
- Volterra, V. 1926. Fluctuations in the abundance of a species considered mathematically. Nature 118:558-560,
- Wächtershäuser, G. 1990. Evolution of the first metabolic cycles. Proceedings of National Academy of Science, 87:200-2004.
- West, G.B., Brown, J.H., Enquist, B.J. 1997. A general model for the origin of allometric scaling laws in biology. Science 276, 122–126.
- Whitfield, J. 2004. Ecology's big, hot idea. PLoS Biol. 2004 December; 2(12): e440.
- Whittaker, R.H. 1965. Dominance and diversity in land plant communities. Science 147:250-260.
- Whittaker, R.J. 1998. Island biogeography. Oxford University Press, Oxford, 285 pages.
- Wiens, J. A. 1984. On understanding a non-equilibrium world: myth and reality in community patterns and processes. In D. R. Strong, D. Simberloff, L. G. Abele, and A. B. Thistle (eds.) *Ecological Communities: Conceptual Issues and the Evidence*, Princeton University Press, Princeton, pp. 439-457.
- Williams, G. C. 1966. *Adaptation and Natural Selection: A critique of some current evolutionary thought*. Princeton University Press, Princeton.
- Wilson, D.S. 1988. Holism and reductionism in evolutionary ecology. Oikos, 53:269-273.

Yodzis, P. 1989. Introduction to Theoretical Ecology. Harper & Row, 384 pages.

Date	Concept or	Seminal paper	Comments on influence
	theory		
1838	Population	Verhulst 1838	Spawned substantial theoretical and empirical
	growth model		work
1880	Food web	Camerano 1880	Early attempts at capturing connections among
	models		species from the equilibrium perspective
1887	The ecological	Forbes 1887	Articulated a persistent theme of nature's
	system as		balancing act and oneness
	microcosm		
1899	Ecological	Cowles 1899	Spawned substantial empirical and conceptual
	succession		work
1926	Competition	Volterra 1926	Laid foundations for a strong and vibrant field
	theory		of theoretical and empirical studies on
	(predation		interspecific interactions in the context of
	models		limiting resources, habitat variability, structured
	included)		populations, time lags, heterogeneity, and many
			others
1934	The niche	Gause 1934	As a non-technical (pre-theoretical concept),
			niche was used earlier (e.g., Elton 1927).
			However, Gause's paper spawned a mix of
			substantial empirical and theoretical work,
			primarily on competition but not exclusively
1935	The climax	Tansley 1935	Although originally introduced by Cowles
			(1899), climax concept was refined by Tansley
			(1935). Subsequently, the concept affected
			greatly debates about the nature of change in
			composition of ecological systems, primarily
			plant associations

Table 2.1. Major theoretical developments, their approximate appearance, and comments on their influence.

1942	Energy pyramid	Lindeman 1942	Led or interacted with numerous ecological ideas, from foraging theory to metabolic theory of ecology; both in theoretical and empirical realms
			icamis
1943	Models of species abundance	Fisher, Corbet, Williams 1943	Spawned voluminous amount of theoretical and empirical work
1948	Leslie's matrices and life tables	Leslie 1948	Provided a foundation for a tremendous amount of refinement to analysis of population dynamics
1951	Dispersal (diffusion models)	Skellam 1951	Boosted primarily theoretical work in the area of foraging theory
1953	Ecosystem	Odum and Odum 1953	Spawned primarily empirical work cast in terms of physical constraints (the ecosystem concept itself was formulated 20 years earlier)
1956*	Functional response models	Holling 1959	Provided a major building block for optimum foraging theory; spawned a mix of empirical and theoretical studies
1963	Equilibrium island biogeography	MacArthur and Wilson, 1963	Became one of the main engines of ecological sciences, with largely empirical but also some theoretical work
1966	Optimal foraging theory	MacArthur and Pianka 1966	Spawned abundant theoretical and empirical research, much with links to evolution
1969	Metapopulation theory	Levins 1969	Provided formal tools for investigation of population dynamics of fragmented habitats and, ultimately, contributed to the rise of metacommunity framework

1969	Individual	Porter and Gates	Provided a clear framework for ecophysiology
	thermodynamic	1969	and evolutionary ecology
	budgets		
1969	Theory of	Odum 1969	Refinement of the ecosystem concept
	ecosystem		
	development		
	(homeostasis)		
1982	Hierarchy	Allen and Starr	Brought awareness of scale to the forefront of
	theory	1982	data interpretation; suggested cross-scale
			analyses
1997	Neutral theory	Hubbell 1997	Spawned considerable amount of empirical
	of diversity		(testing) and theoretical (refinement) research;
			long-term impact still unclear

* - paper presented in 1956 but published in 1959.



Figure 2.1. Haeckel's depiction of mosses. Note his desire to place various species together as a natural system. Kunstformen, 1904 - plate 72.



Figure 2. 2. Chronology of major theories and concepts. The y axis represents a particular concept. Choice of the initial dates was based on either a foundational paper for a given theory (e.g., Lindeman's (1942) analysis of lake energy flows or Hubbell's (1997) neutral theory of diversity), on an early influential paper (e.g., Forbes' microcosm perspective), or on a paper that revived or reinvigorated an already existing and accepted idea (e.g., Odum's 1953 perspective on ecosystem concept). The only motivation in each case was to anchor a theory or an important concept in a time period indicative of the beginning of its rise to prominence. Even, if one reshuffled the order of appearance to satisfy some other, equally valid criteria, the overall temporal trend would not change. The selection of the theoretical components, with the exception of the three most recent, was guided by the compilation "Foundations of Ecology" by Real and Brown (1991) and by characteristics of propositions as introduced in Chapter 1.



Figure 2.3. Holdridge vegetation types are defined by mean annual 'biotemperature' (growing season and temperature index), annual precipitation, and a ratio of potential evapotranspiration to mean total annual precipitation.



Figure 2.4. Number of papers found after searching for "competition" (deemed as a reference term with a balance of empirical and theoretical connotations), "hierarchy theory + scale" and "island biogeography" by JStor search engine between 1900 (for competition) and 1930 (for Island Biogeography, IBT, and Hierarchy) and the year 2000. The number of all three categories of papers increased exponentially but their slopes were higher for the theory-based papers (0.85 and 1.18 for Hierarchy and IBT, respectively) as compared to competition (0.67).



Figure 2.5. A simplified way of arranging ecological models relative to each other in order to assess the strengths and gaps. It is possible to place models along 3 axes of typical tradeoffs of realism, generality, and precision, together with some examples. Partial overlaps of areas suggest possible models of an intermediate or mix nature.

Chapter 3: A General, Unifying Theory of Ecology? Jay Odenbaugh

Samuel Scheiner and Michael Willing (Chapter 1) have provided a philosophical framework for understanding ecological principles, theories, and models. Fundamentally, they contend that contrary to many ecologist's views about their own discipline, ecology already possesses a general, unified theory. In this essay, I first present their framework. Second, by way of comparison, I consider the work of the population and community ecologist Robert MacArthur. MacArthur's own work was thought of as providing unifying theories though I argue it focused more specifically on integrative theories and models. Finally, I expand on several points in the Scheiner and Willig framework.

THE SCHEINER-WILLIG FRAMEWORK

According to Scheiner and Willig (SW), a theory in the sciences consists in two elements, a set of principles and a domain (Chapter 1, pp. 2 - 4). What is the domain of ecology? According to SW, it is the spatial and temporal patterns of the distribution and abundance of organisms which includes the causes and consequences of ecological processes (Chapter 1, p. 6). Lest one think this is too "organism-focused", it includes biotic and abiotic factors which affect organisms along with groups of organisms at a variety of levels including populations, communities, ecosystems, and landscapes. The principles describing this domain are selected on the basis of two criteria (2008, 23):

- Inclusionary rule: For something to be a principle of a given domain, the principle must be shared by many constituent theories of that domain.
- Exclusionary rule: For something to be a principle of a given domain, it must distinguish this domain from some other distinct domain.

Put simply then,

Something is a principle of a given domain just in case it is shared by the constituent theories of that domain and is not shared with constituent theories of a distinct domain. So, what are the basic principles of this general and unified theory of ecology? They are principles that ecologists already accept given the domains they study. In a way, SW are simply

making explicit what those principles are. More importantly, they are presenting them together which highlights their role in structuring ecological thought collectively. Though ecology may appear to be a fairly disunified discipline, there is actually much unity undergirding various theories and models. Here is their list of these unifying principles (Chapter 1, p. 18).

64

- 1. Organisms are distributed in space and time in a heterogeneous manner.
- 2. Organisms interact with their abiotic and biotic environments.
- Variation in the characteristics of organisms results in heterogeneity of ecological patterns and processes.
- 4. The distributions of organisms and their interactions depend on contingencies.
- 5. Environmental conditions are heterogeneous in space and time.
- 6. Resources are finite and heterogeneous in space and time.
- Birth rates and death rates are a consequence of interactions with the abiotic and biotic environment.
- 8. The ecological properties of species are the result of evolution.

SW recognize that theories come in degrees of generality or abstractness. They distinguish between general theory, constituent theory, and models (Chapter 1, p. 4). A general theory consists in confirmed generalizations which are abstracted from facts that have been systematically tested. Using general theory in addition to more concrete considerations, we can arrive at constituent theories. A constituent theory consists in confirmed generalizations or laws and from them models may be derived. Finally, there are models which are propositions by which hypotheses can be articulated and evaluated. Thus, there is a hierarchy of theoretical structures which become less abstract as one moves down the hierarchy. The principles of the general theory are listed above and the constituent theories include succession theory, foraging theory, metapopulation theory, and many others.

As a particular example of the SW framework consider work on predator-prey theory. There are several general principles which can be used to derive constituent theory concerning predator-prey relationships which include the claim that organisms interact with their abiotic and biotic environments, resources are finite and heterogeneous in space and time, and birth rates and death rates are a consequence of interactions with the abiotic and biotic environment. Let's suppose we assume that the growth rate of the prey is determined by the growth rate of the prey population independent of the predator minus the capture rate of prey per predator multiplied by the number of predators and the growth rate of the predator population is determined by the rate at which each predator converts captured prey into predator births minus the rate at which predators die in the absence of prey multiplied by the number of predators. Here we have a classical theory of predator-prey interactions. From this we can derive more concrete models which may be tested by data. For example, let's further assume as Lotka-Volterra models do that the prey grows exponentially in the absence of the prey, predator and prey encounter one another randomly in proportion to the abundance, the predators have a linear functional response, the numerical response of the predator is a constant multiplied by the functional response, and the predator declines exponentially in the absence of the prey. Here we start with general principles, add detail which results in a constituent theory and then only arrive at a model once we have added in this case quantitative detail. It is crucial to note that from general principles one can derive many different constituent theories and from them many different models. We could have devised a different model by assuming a different type of functional response for example.

Each of the principles above can be found in different areas of ecology though articulated in different ways. For example, consider principle (1) – organisms are distributed in space and time in a heterogeneous manner. In population ecology, organisms are distributed unevenly as a population over habitat. For example, they may be distributed unevenly vertically in a lake or in a forest. Likewise, in metapopulation theory, organisms are grouped into a population of populations and their dynamics are largely controlled by local extinction and migration amongst distinct patches. This may be the case with forest patches or oceanic islands. Similarly, in landscape ecology, we see organisms distributed in different ecosystems or biomes. Depending on the sub-discipline of interest, each of these principles can be made concrete in different ways. As another example, consider the general principle (3) – variation in the characteristics of organisms results in heterogeneity of ecological patterns and processes. In behavioral ecology, the characteristics of interest involve different foraging strategies such as being generalist or specialist and in life history theory it may involve the characteristics of being an annual or perennial plant. Finally, consider the principle that the distributions of organisms and their interactions depend on contingencies. We can see how this principle is made concrete in different ways by considering the introduction of stochastic growth rates in population ecology and in community ecology through the notion of "ecological drift" in neutral theories of biodiversity.

The fact that ecological theory consists in this hierarchy of general principles, constituent theories, and models has fundamentally important implications for evaluating general principles, constituent theories, and models. Specifically, it challenges an excessively narrow Popperian view of theory testing in ecology. Suppose a model fails some specific test or tests. This does not necessarily impugn the constituent model from which it was derived nor the general principles on which it is based. The general principles and constituent theories are made concrete in models based on the specific assumptions, backgrounds, and definitions which are considered. Thus, one falsifies a general principle only after many different constituent theories and models have been evaluated in light of the relevant facts.

SW have focused on the notion of generality and unification as the relations that exist between ecological theories and data about which more will be said below. However, I now will introduce what I believe is a different conception of how theories relate, namely that of integration. As an example, let's consider the work offered by the eminent ecologist Robert MacArthur.

MACARTHUR'S INTEGRATIVE APPROACH

Robert MacArthur stands as one of the most influential and controversial ecologists ever to work in the discipline (Fretwell 1975, Pianka and Horn 2005). He is recognized for having done exceptionally original theoretical and empirical work. However, many believe that he took ecology down the wrong path both theoretically and methodologically. As an example, MacArthur and his colleagues' work on limiting similarity is often seen in this light (MacArthur and Levins 1967, May and MacArthur 1972). The project was to understand why species are spaced along resource spectra given their niche breadths and widths respectively. The MacArthurites argued that we should expect that the ratio of niche breadth and width should be approximately one. However, this theoretical result was argued to be very fragile on the assumptions made in the models and did not hold up with respect to the data (Abrams 1983). So, some believe it was a dead end in theoretical ecology. Independent of one's opinion on this matter, MacArthur's work serves as an interesting case study. He too attempted to provide a framework for understanding how ecological theories relate to each other. However, in my mind, it was not a unificationist but an integrationist approach. More on this later; first, a bit of history (for more details, see Odenbaugh (2006)).

Robert MacArthur, E. O. Wilson, Egbert Leigh, Richard Levins, Leigh van Valen, and Richard Lewontin met at MacArthur's lakeside home in Marlboro, Vermont in July 1964. The subject of their conversation was their own research in evolutionary genetics, ecology, and biogeography and the overall future of what is termed "population biology". Ironically, the subject matter of these conversations was not simply population biology understood as including population genetics, population ecology, and possibly ethology. It clearly included disciplines like community ecology, macroevolution, and biogeography given the sorts of models formulated and questions asked. More importantly, there was a general tendency to approach these areas with mathematical theory best represented in theoretical population biology. For two days, each participant discussed their work and how a "central theory" could be achieved (Wilson 1993, 252-253). The work that resulted from these collaborations was important and changed much of evolutionary and ecological theory. Collaborator E. O. Wilson and mentor G. E. Hutchinson wrote the following of MacArthur after his death in 1972.

[He] will be remembered as one of the founders of evolutionary ecology. It is his distinction to have brought population and community ecology within the reach of genetics. By reformulating many of the parameters of ecology, biogeography, and genetics into a common framework of fundamental theory, MacArthur – more than any other person who worked during the decisive decade of the 1960s – set the stage for the unification of population biology. (1982, 319)

Did MacArthur and his co-workers "unify" evolution, ecology, and biogeography? I will argue contrary to Wilson and Hutchinson he did not.

In order to assess MacArthur's accomplishment, we must understand the components of the program he and others articulated. Here are some of the elements. First, MacArthur typically formulated general, simple deterministic models which lacked precision. In the terms of Richard Levins' account of model building (1966), precision was sacrificed for generality and realism. This is not to say that MacArthur modeled ecological systems realistically; rather, the desiderata of interest were generality and realism and precision less so. As an example MacArthur's "realism", he devised a mechanistic consumer-resource model with two consumers and two resources and showed how the more phenomenological Lotka-Volterra interspecific competition could be derived from it (MacArthur 1972). Second, MacArthur also emphasized the ecological process of interspecific competition as a mechanism structuring ecological communities. This is

evident is his work on limiting similarity and species distributions (i.e., the "broken stick" model). This is not to say that he did not work on other types of processes like predation (MacArthur 1955); rather it is that interspecific competition played a predominate role in his thinking. Third, MacArthur rarely evaluated model predictions statistically. There are of course exceptions to this rule but mostly he and his colleagues evaluated their models by looking for corresponding dynamical patterns such as stable equilibria and various types of cycles. Finally, he was a master at presenting complex mathematical results with graphical representations (MacArthur and Wilson 1967, 1970). Specifically, MacArthur used isocline analysis to not only present theory in pedagogically useful ways but also to draw interesting and unobvious implications (Rosenzweig and MacArthur 1963).

MacArthur and his colleagues produced a variety of different models involving environmental heterogeneity, density-dependent selection, optimal foraging, limiting similarity, and equilibrium island biogeography. As an example of MacArthur's theoretical work, let us consider his modeling of density-dependent selection. This is a case where MacArthur attempts to integrate ecological and evolutionary concepts which connects to several of SW's principles (specifically principles (2), (5), and (8)).

In most evolutionary models according to MacArthur, population geneticists use r, the intrinsic rate of increase of a population, as a measure of fitness. He writes,

For populations expanding with constant birth and death rates, r, or some equivalent measure (Fisher used r; Haldane and Wright used e^r which Wright called W) is then an appropriate definition of fitness (1962, 146).

However, as MacArthur notes, present values of r may not be reliable predictors of the number of descendants a group of individuals will have since r is an accurate measure of fitness only if the environment is relatively stable. One way in which the environment may be unstable is if population density affects fitness. In fact, MacArthur writes, "[t]o the ecologist, the most natural way to define fitness in a crowded population is by the carrying capacity of the environment, K,...." (1962, 146). MacArthur offers the following mathematical model. Let n_1 and n_2 represent populations of alleles 1 and 2 respectively and let them be governed by the following equations

$$dn_1/dt = f(n_1, n_2)$$
(1)
$$dn_2/dt = g(n_1, n_2)$$
(2)

To understand this model, it is simplest to examine it graphically (Figure 3.1).

Suppose we have a phase space where the *x*-axis represents the population n_1 of allele 1 and the other *y*-axis represents the population n_2 of allele 2. Thus, a point in the space represents the joint abundances of population n_1 and n_2 . Let us suppose there is a set of values of n_1 and n_2 such that there is a solution $f(n_1, n_2) = 0$, or equivalently, $dn_1/dt = 0$ for those values of n_1 and n_2 . If the population of n_1 is to the left of the *f*-isocline, then it will increase. Likewise, if the population of n_1 is to the right of the *f*-isocline, then it will decrease. Let us further suppose that there are a set of values of n_1 and n_2 such that there is a solution $g(n_1, n_2) = 0$, or equivalently, $dn_2/dt = 0$ for those values of n_1 and n_2 . If the population of n_2 is below the *g*-isocline, then it will increase. Likewise, if the n_2 population is above the *g*-isocline, then it will decrease.

There are four different ways the two isoclines can relate to one another. In part A of the Figure 3.1, we can see from the vector arrows that allele 1 will outcompete allele 2. Likewise, in part B, allele 2 will outcompete allele 1. In part C, the vector arrows show that there is a stable equilibrium between allele 1 and 2. Finally, in part D, the alleles whichever allele is more frequent at the outset will outcompete the other.

We can now explain how this model represents both ecological and evolutionary features. The *f*-isocline intersects the n_1 coordinate at the value K_{11} . In this circumstance, the population consists only of allele 1 and K_{11} represents the number of allele 1 homozygotes that can maintain themselves in this environment. In other words, K_{11} is the carrying capacity of the allele 1 homozygotes. Likewise, the *f*-isocline intersects the n_2 coordinate at the value K_{12} . K_{12} is the number of allele 2 which can keep allele 1 from increasing and represents the carrying capacity of the environment for heterozygotes expressed in units of allele 1. We can similarly designate the end of points of the *g*-isocline as K_{22} and K_{21} . MacArthur concludes, "We have now replaced the classical population genetics of expanding populations, where fitness was *r*, as measured in an uncrowded environment, by an analogous population genetics of crowded populations where fitness is *K*" (1962, 149).

Let us now consider what in fact MacArthur accomplished theoretically in this and other examples. First, let me define the notion of a unifying theory.

A unifying theory applies a single theoretical framework (for example, common state variables and parameters) to a variety of different phenomena.

Often philosophers of science consider a theory to be a *unifying theory* in just the sense I defined above (see Friedman 1974, Kitcher 1989, Morrison 2000 for discussion and debate). As Margaret Morrison writes of Newtonian mechanics and Maxwell's electrodynamics,

The feature common to both is that each encompasses phenomena from different domains under the umbrella of a single overarching theory. Theories that do this are typically thought to have "unifying power"; they unify, under a single framework, laws, phenomena or classes of facts originally thought to be theoretically independent of one another (2000, 2).

In the case of Newtonian mechanics, Newton showed how his laws of motion and the law of universal gravitation could account for both Galileo's laws of terrestrial mechanics and Kepler's laws of planetary motion.

One might argue that the model of density-dependent selection just provided shows that MacArthur actually unified parts of population and community ecology. The equations describing the density-dependent apply likewise to competing species in the familiar Lotka-Volterra model of interspecific competition. Put differently, MacArthur applied exactly the same framework to different genotypes within a population that he applied elsewhere to different populations within a community. Just as Newton showed that the same laws apply to the sublunary domain as to the supra-lunary domain, MacArthur showed the same biological principles apply to genotypes and species. However, I would resist this conclusion since the two cases are importantly disanalogous. In the Newtonian case, we recognize that sub-lunary and supra-lunary domains actually are one domain and hence any principle which applies to the former should apply to the latter. However, in MacArthur's case, we have two distinct domains, alleles and species. An analogous case would be between work done using diffusion equations in physics and spatial ecology. Both use the same mathematical formalism but apply them to different domains altogether. We would not say the same laws apply in both cases; rather both use similar mathematical tools.

If unification with respect to scientific theories or models minimally consists in a "single overarching theory" accounting for a variety of phenomena, then it appears that MacArthur's framework could not have unified population biology. If one examines the various models that MacArthur devised – models of environmental heterogeneity, density-dependent selection, limiting similarity and equilibrium biogeography for example – they form an extremely diverse

70

group. The state variables and parameters are rarely the same across models; i.e., they are rarely even representing the same phenomena. The state variables of the limiting similarity models are population abundances and the parameters are intrinsic rates of growth, carrying capacities, and interaction coefficients whereas in the equilibrium models of island biogeography the state variable is species richness and the parameters are rates of immigration and extinction. Likewise, in the density-dependent selection model presented above, the state variables are populations of alleles and the parameters are carrying capacities. There is no common overarching structural framework to unify population genetics, population and community ecology, and biogeography. This is not to say that there is nothing that these models have in common of course. However, the common ingredients are usually that the models represent equilibrium behavior, make important optimality assumptions, and are represented with deterministic equations. Nonetheless, that which is at equilibrium is sometimes population abundances, species numbers, or population of alleles. Likewise, what is considered optimal is sometime phenotypes, genes, or the numbers of species and their abundances in a community. This methodology certainly does not generate a theory like Newtonian mechanics which consists in a small set of schematic equations concerning the motion of objects. Hence, MacArthur provided no theoretical framework of the sort needed to unify population biology.

Nonetheless, MacArthur did show how one could represent *both* evolutionary and ecological factors at different scales in mathematical models. These different areas of population biology had largely proceeded independently of one another. However, if evolutionary and ecological processes are commensurate, then it was increasingly important to theoretically integrate these different processes at work in biological systems. It surely is correct that MacArthur "brought population and community ecology within the reaches of genetics" as claimed by Wilson and Hutchinson. However, he did not do so by "reformulating many of the parameters of ecology, biogeography, and genetics into a common framework of fundamental theory". We can now see how MacArthur approached the relations between theories. Here is another definition.

An integrating theory takes a variety of theories (different state variables and parameters) and combines them in their application to a variety of phenomena.

He supplied a variety of models that incorporated many different evolutionary and ecological state variables and parameters thus taking a first step toward integrating population biology.

The SW program differs from the MacArthur program in that it explicitly looks for *common* principles across ecology's sub-disciplines whereas the MacArthur program was looking for "piecemeal" connections. In fact, we can "harmonize" the SW and MacArthur programs if we recognize that both are emphasizing different parts of the theoretical hierarchy. SW have worked hard in identifying the key general principles which constituent theories and models share. MacArthur spent most his time attempting to articulate novel constituent theories and more specifically models for understanding the distribution and abundance of organisms. Thus, unification and integration are regions along a continuum of theoretical structures. Unification is largely to be found at the most abstract level of the theoretical spectrum. Integration on the other hand is to be found at the level of constituent theories and models. These two features of theoretical structures are points of emphasis and are compatible with one another.

Having said all of this, MacArthur at times was clearly engaged in the same project as SW. MacArthur famously wrote,

Science should be general in its principles. A well-known ecologist remarked that any pattern visible in my birds but not in his *Paramecium* would not be interesting, because, I presume, he felt it would not be general. The theme running through this book is that the structure of the environment, the morphology of the species, the economics of species behavior, and the dynamics of population changes are the four essential ingredients of all interesting biogeographic patterns. Any good generalization will be likely to build in all these ingredients, and a bird pattern would only be expected to look like that of *Paramecium* if birds and *Paramecium* had the same morphology, economics, and

dynamics, and found themselves in environments of the same structure (1972, 1) Clearly, MacArthur believed that there were general principles concerning morphology, economics and dynamics which would be made concrete in possibly different ways in different constituent theories and models depending on the taxonomic group under study.

ELABORATING ON THE SW PROGRAM

Let us now turn again to the SW program and consider a few questions. First, how do we distinguish between general theory, constituent theory, and models? Are they different in kind or degree? One might argue that they are not different in kind but vary continuously along some dimension. For example, a principle is more general than another when the former's domain is a
superset of the former's domain; or conversely a domain is a proper subset of another. However, one might also argue that structures differ in kind. For example, it is customary to believe theories are composed of a small set of natural laws. Consider Newtonian mechanics with its three laws of motion and gravitation as familiar case in point. Models, on the other hand, are often thought of as not consisting in natural laws at all; rather, they are idealized representations of natural systems without natural laws. If this is so, then theories and models are distinct. So, there is a general question about how these different sorts of structures relate to one another.

Second, biologists make much out of the notion of *contingency* and SW do so in (3). However, what is "contingency"? In what sense it is a cause of the ecological patterns? Here is one way of construing contingency.

An effect variable Y is *contingent* on a causal variable X to the degree that slightly changes in values of X subsequently change values of Y greatly.

Of course, this is just sensitivity to initial conditions - a species of nonlinearity - and there are various quantitative measures of it. Moreover, we could generalize with regard to a multivariable system where small changes in a set of causal variables $X_1, X_2, X_3, \ldots, X_n$ lead a large change in the effect variable Y. In the way that I have characterized contingency, it is not a cause of anything; rather, it is a pattern concerning causes and their effects. There may be other characterizations that render it a cause amongst others.

Third, SW claims that evolution causes the ecological properties of species. As the eminent ecologist G. E. Hutchinson argued ecology is the theater of the evolutionary play. Put less metaphorically and only in terms of natural selection, ecological processes create selective regimes. These ecological processes cause or determine mechanistic or proximate differences in reproductive success. One way of construing SW's insight is that they are insisting that current ecological processes are in play because of *past* evolutionary processes. Hutchinson's idea can then be coupled to this proposition with the claim that *current* evolutionary processes are in play because of *past* ecological processes. Thus, ecological and evolutionary processes are spatiotemporally interdependent. Thus, there is a crucial interaction between ecological and evolutionary processes. By emphasizing this point, ecology and evolution can be integrated even further.

Finally, where does ecosystem ecology fit in the prescribed domain of the abundance and distribution of organisms? The domain of ecosystem ecology is roughly the cycling of nutrients

and flow of energy. For example, ecosystem ecologists focus on the nitrogen and carbon cycles or gross and net primary production. One could and some do argue that ecosystem ecology really just is biogeochemistry and not ecology per se since organisms – the currency of ecology and other biological sciences – have disappeared from the science (Cooper 2003). However, in my view, this would inject a bias in favor of population and community ecology and the history of ecology has been ensconced with ecosystem ecology just as much as these other disciplines. In fact, historians of ecology have spent more time writing about ecosystem ecology than population or community ecology (Golley 1996, Hagen 1992). I am unsure of why this is but it is an interesting fact about the history, or historians, of science.

If ecosystem ecology is a genuine branch of ecology as I have suggested that it is, then SW implicitly subsume this domain under those principles. In my view, this is where a integrative framework is important since it can couple energy flows and nutrient cycles with food web dynamics for example. Of course, there may be even more general principles one can provide which brings ecosystem, community, and population ecology together.

CONCLUSION

In this essay, I have presented the SW unification framework and have also presented a similar though importantly different integrative framework through the work of Robert MacArthur. Importantly, unification concerns finding the most general principles of a domain and integration consists in bringing together different constituent theories and models. However, unification and integration are complementary since they concern different regions of the theoretical hierarchy. I also considered some specific elements of the SW framework including the notion of contingency, the relationship between ecology and evolution, and the place of ecosystem ecology in their general principles. Whether SW have provided a complete account of the unifying principles of ecology, they have certainly made an excellent and productive start.

Literature Cited

- Abrams, P. (1983) "The Theory of Limiting Similarity" *Annual Review of Ecology and Systematics* 4: 359 – 76.
- Cooper, G. (2003) *The Science of the Struggle for Existence: On the Foundations of Ecology*, Cambridge University Press.
- Fretwell, S. (1975) "The Impact of Robert MacArthur on Ecology," *Annual Review of Ecology* and Systematics 6:1-13
- Friedman. M. (1974) "Explanation and Scientific Understanding," *The Journal of Philosophy*, Vol. 71, No. 1.: 5 – 19.
- Golley, F. (1996) A History of the Ecosystem Concept: More than the Sum of its Parts. Yale University Press.
- Hagen, J. (1992) An Entangled Bank: The Origins of Ecosystem Ecology. University of Chicago Press.
- Kitcher, P. (1989) "Explanatory Unification and the Causal Structure of the World," in P.
 Kitcher and W. Salmon (eds.), *Scientific Explanation. Minnesota Studies in the Philosophy of Science*, vol. 13. Minneapolis: University of Minnesota Press: 410-505.
- Levins, R. (1966) "The Strategy of Model Building in Population Biology," *American Scientist* 54: 421-431.
- MacArthur, R. (1955) "1955. Fluctuations of animal populations, and a measure of community stability," *Ecology* 36: 533-536.
 - _____ (1962) "Some generalized theorems of natural selection," *Proceeding of the National Academy of Science*, USA. 48.
 - (1970) "Graphical analysis of ecological systems," 61-72 in *Some Mathematical Questions in Biology*, American Mathematical Society.

(1972) *Geographical Ecology*. Princeton University Press.

- MacArthur, R. and R. Levins (1967) "The limiting similarity, convergence and divergence of coexisting species," *American Naturalist* 101: 377-385,
- MacArthur, R. and E. O. Wilson (1967) *The Theory of Island Biogeography*. Princeton University Press.
- R. May and R. MacArthur (1972) "Niche overlap as a function of environmental variability," *Proc. Nat. Acad. Sci. USA* 69: 1109-1113.

- Morrison, M. (1000) Unifying Scientific Theories: Physical Concepts and Mathematical Structures. Cambridge University Press.
- Odenbaugh, J. (2006) "The Strategy of 'The Strategy of Modeling Building'", *Biology and Philosophy*, 607-621.
- Pianka, E. and H. Horn (2005) "Ecology's Legacy from Robert MacArthur,". Chapter 11 (pp. 213-232) in K. Cuddington and B. Biesner, eds. *Ecological Paradigms Lost: Routes of Theory Change*, Elsevier Academic Press, 213 232.
- Rosenzweig, M. and R. MacArthur (1963) "Graphical representation of stability conditions of predator-prey interactions. *American Naturalist* 97: 209-223.
- Scheiner, S. and M. Willig (2008) "A General Theory of Ecology," *Theoretical Ecology* 1: 21 28.
- Wilson, E. O. (1993) Naturalist. Harper Collins.
- Wilson, E. O. and G. E. Hutchinson (1982) "Robert Helmer MacArthur 1930 1972, A Biographical Memoir," *National Academy of the Sciences*.



Figure 3.1. Density-dependent selection and competitive exclusion of alleles (MacArthur and Wilson 1967, 147).

Chapter 4: Foraging Theory Andrew Sih

This chapter is about the predator-prey interaction viewed at the individual level. It discusses ecological theory designed to explain or even predict how predators and prey adjust their behavior in response to changes in their external and internal environments. The resulting patterns of predation often then have major impacts on population dynamics, species interactions (e.g., competition and predator-prey interactions), community dynamics, and ecosystem patterns. As such, foraging theory links with and potentially provides a mechanistic underpinning for many of the other constitutive theories in ecology.

The underlying framework for foraging theory, as well as for evolutionary ecology, in general, is the optimality approach. The key premise is that natural selection has shaped organisms to exhibit foraging behaviors that enhance lifetime fitness. Behaving optimally is a complex endeavor. It requires the organism to account for its own relatively fixed traits (e.g., its morphology, physiology), its current state (e.g., energy reserves), its biotic environment (e.g., competitors, predators, diseases) and abiotic environment. Often, organisms have only imprecise information about many, if not all of these factors. While we do not expect organisms to necessarily be capable of coming up with the optimal behavior balancing this complicated mix of factors, the notion is that the optimality framework can predict qualitative behavior well enough to be useful.

Here, I will: 1) define the domain of foraging theory – the fundamental questions and research goals; 2) outline basic propositions that guide the theory; 3) describe propositions on classes of factors that influence foraging behavior; 4) outline several propositions on general predictions (outcomes) foraging theory; 5) provide a brief overview of major types of models including both early simple models and more complex models that include more aspects of reality; 6) discuss links to other fields of ecology; and 7) note promising future directions.

DOMAIN

In keeping with the notion that "the domain of ecology and its general theory is the spatial and temporal patterns of the distribution and abundance of organisms" (Scheiner and Willig 2005), the domain of foraging theory is the spatial and temporal patterns of behavior of organisms

foraging for resources. Although predation, 'tooth and claw', might conjure up an image of carnivores chasing down animal prey (e.g., lions chasing ungulates, big fish chasing smaller fish), foraging theory can be applied to a very broad range of consumers including: predators foraging for prey, herbivores foraging on plants, parasites foraging for hosts, pollinators foraging for nectar or pollen, frugivores searching for fruit, breeding adults searching for oviposition sites, and even plants sending out roots or shoots in search of resources. Although this may be taking some liberty with semantic fine distinctions, I will interchangeably use the terms predators, consumers or foragers foraging on prey or resources to refer to any of these systems. When three trophic levels are involved, I will refer to them as predators consuming foragers that consume prey. Regardless of the type of forager, key goals are to explain the forager's allocation of time or energy to alternative foraging options, e.g., habitats, patches, times of day, or food types.

Foraging theory has found it useful to partition the overall foraging process into 2 main stages: 1) behaviors that influence encounter rates between predators and prey; and 2) behaviors that influence the probability of consumption given an encounter (Figure 4.1). Stage 1 involves decisions by predators on where to forage (habitat or patch use), when to forage (diurnal or seasonal activity patterns), and details of foraging mode (active vs. sit-and-wait foraging, foraging speed, search path etc). The main focus of foraging theory for this stage has been on where to forage – optimal patch use (Stephens and Krebs 1986). Stage 2 is often split into several steps: the probability of consumption given prey capture. The main focus of foraging theory in this stage has been on the forager's decision on whether to attack prey that are encountered– i.e., on predicting the prey types that should be part of the forager's optimal diet (Stephens and Krebs 1986; Sih and Christensen 2001).

Patch use and diets have classically been viewed as key parts of an organism's niche (Schoener 1971, 1974, 1989; MacArthur 1972; Chase Chapter 4). Thus, in a sense, the domain of foraging theory is to explain major parts of individual and species niches in functional terms relating to foraging and avoidance of predation.

BASIC PROPOSITIONS

The basic premise of foraging theory is that foraging patterns reflect behavioral rules that have evolved under natural selection to be adaptive within constraints. This premise obviously relates

to the Theory of Ecology's fundamental principle number 8 – the evolutionary cause of ecological properties. If foraging behaviors have indeed been shaped by natural selection, then the adaptive evolutionary ecology framework, often referred to as the optimality approach, should be insightful. The optimality approach as applied to foraging involves 3 basic propositions (Table 4.1).

80

Proposition 1

Foraging patterns maximize fitness or a correlate of fitness. In the optimality approach, the investigator must first specify an objective function or goal. Since natural selection maximizes fitness, in evolutionary ecology theory, the ideal goal is fitness maximization. However, because fitness is often difficult to measure, much of foraging theory is based on the simpler, surrogate goal of maximizing net energy intake (Stephens and Krebs 1986; Stephens et al. 2007). The idea is that having more energy available often translates into higher fitness (better long-term survival, greater reproductive success). While other considerations (e.g., mating, avoiding risk) clearly can influence fitness, perhaps during foraging bouts, it is often reasonable to assume that net energy intake rates are positively correlated with fitness. The underlying assumptions that resources are limiting, and that adaptive foraging behavior is required to enhance resource intake relate to fundamental principle 6 (finite and heterogeneous resources).

Proposition 2

Foraging patterns depend on the range of options available to the forager and on how each available option affects fitness or a correlate of fitness. This proposition is where the biologist's expertise comes into play. Applying propositions 2 requires insight on constraints that limit available options (e.g., physicochemical, genetic, physiological, or morphological constraints) and on the mechanisms for how ecological and social factors, along with constraints might influence the relationship between behavior and fitness. For example, if we ask how deep a duck should dive to forage and what prey should it take, a good answer probably requires a complex mix of considerations involving energy availability in different habitats with different prey, energy costs, oxygen loads and perhaps temperature and light considerations. Most of these factors depend on the duck's physiology and energy state. Furthermore, the duck's optimal foraging behavior could also depend on competition with other ducks and with other animals foraging on the same prey, and on predation risk. In that case, one might want to include information about the traits of potential competitors and predators. Foraging theory, at its best, accurately includes all factors that need to be included to understand the forager's behavior.

Proposition 3

Foraging behavior balances conflicting demands – tradeoffs are important in shaping foraging behavior. Tradeoffs are at the heart of many ecological patterns. Indeed, one could say that in the absence of tradeoffs we would see much less organismal diversity (see Chase Chapter 4; Fox et al. Chapter 14). Without tradeoffs, the one type of organism (in each trophic level) that exhibits superior performance in all situations would dominate, outperforming all other types. In reality, of course, due to tradeoffs, no one type can do best in all environments. A species that is a great competitor in one environment typically does poorly in other environments (Tilman 1982). Even in any given environment, the best competitor often does not cope well with predators or other enemies (Sih 1987; Werner and Anholt 1993).

In many cases, the mechanism underlying performance tradeoffs involves foraging behavior. Foragers that exhibit superior foraging performance in one environment typically do not forage as well in other environments, and foraging behaviors that increase net energy intake rates (e.g., high activity, use of habitats with more resources) often also expose foragers to high predation risk (Sih 1980, 1987; Werner and Anholt 1993; Lima 1998; Bednekoff 2007). The flip side of this is that foraging behavior is often influenced by multiple conflicting demands; i.e., predation risk and other fitness-related factors often have important effects on foraging decisions.

With the biology of the system well represented, the final step in foraging theory involves 'doing the math' – using analytical or computer-based optimization methods to find the option that yields the highest fitness. To emphasize, the key challenge for foraging theory (and for optimality theory, in general) is not the math. Instead, the main challenge is to capture the key elements of the biology of the system associated with propositions 2 and 3.

PROPOSITIONS ABOUT KEY FACTORS INFLUENCING FORAGING BEHAVIOR Proposition 4.

Foraging behavior varies depending on the forager's other traits. This proposition and the following one emphasize the key, simple insight that foraging behavior is variable – among species, among individuals within a species, and within-individuals. It depends on both the individual's traits (e.g., its morphology, physiology, energy state, life history state etc) and on the environmental context (see the next proposition). This variation in foraging behavior relates to fundamental principles 1 (the heterogeneous distribution of organisms), 3 (variation in organismal characteristics affect ecological patterns and processes), 4 (contingency), and 5 (environmental heterogeneity). Foraging behavior depends on the individual's morphology and physiology because the way the animal (or plant) is built clearly influences its ability to search for, detect, recognize, capture and/or consume prey efficiently. In foraging theory, morphology (e.g., limbs, feeding morphology) and sensory physiology (e.g., ability to detect different kinds of prey using various types of cues) are typically considered to be fixed constraints that provide a mechanistic explanation for parameters and functions that determine optimal behavior. For example, a key determinant of the optimal diet is the relative value of different prey. Prey value depends on the prey's assimilatable energy content, the probability of capture, and handling time. Assimilatable energy content depends on the interplay of prey defenses that reduce digestibility and forager digestive physiology. Probability of capture often depends on prey versus predator mobility, which depends on their morphologies. Similarly, handling time often depends on prey and predator morphologies.

Foraging theory predicts how foraging behavior should change in response to changes in the forager's traits. While the morphological and physiological traits discussed in the previous paragraph are usually relatively fixed as compared to behavior per se, nonetheless, they often show plasticity over a lifetime (DeWitt and Scheiner 2004). As a forager grows, its optimal patch use and diet should change. In many cases, the changes in morphology or physiology are induced responses to foraging-related challenges (e.g., induced responses to low food or predation risk while foraging; Tollrian and Harvel 1999; Relyea 2001. Most interestingly, some traits change in direct response to foraging behavior, and in turn, influence current and future foraging behavior. These traits (often referred to as state variables) include the forager's energy reserves, its condition or vigor, and its information state (e.g., how much it knows about the

environment). In a later section, I discuss dynamic models that predict optimal foraging behavior in this dynamic situation.

Proposition 5

Foraging behavior responds to environmental heterogeneity, including other species. Proposition 4 emphasized how foraging behavior responds to the organism's traits, in essence, its internal environment. Obviously, optimal foraging behavior also depends on heterogeneity in the external environment (i.e., variation in resources, competitors, predators (enemies, in general), mutualists and abiotic factors; principles 5 and 6). Much of foraging theory ignores species interactions other than the response of the forager to its resources. That is, most foraging models emphasize forager responses to heterogeneity in the value of the different available prey types (optimal diet theory) and in the spatial distribution of resources (optimal patch use theory). Species interactions, however, obviously also often influence foraging behavior (fundamental principle 7). Accordingly, a branch of foraging theory accounts for how competition and predation alter foraging behavior. Species interactions often involve feedbacks that can make the dynamics quite complex. For example, although prey (while foraging) might shift their patch use in an attempt to avoid predators, predators can then shift their patch use to follow prey (Hugie and Dill 1994; Sih 1998; Luttbeg and Sih 2004; Hammond et al. 2007).

In parallel with the emphasis on scale in other chapters, environmental heterogeneity influences foraging behavior at several spatial scales. Foragers choose a general habitat type in which to forage (e.g., forest vs. meadow). Within that habitat, they choose particular patches, and within each patch, they attack some prey and reject (or ignore) others. Decisions interact across scales. Patch use depends on diet preferences (whether a bird prefers foraging in the forest or in an adjacent meadow can depend on its diet preferences) and patch use can determine diets (if a bird has chosen to forage in the forest, it will only consume forest prey).

Interestingly, although to behavioral ecologists it is obvious that foraging behavior is quite plastic, and that therefore the outcomes (e.g., in terms of predator-prey or competitive dynamics) should be quite plastic, many other branches of ecological theory tend to ignore this plasticity. For example, competition theory often assumes a constant competition coefficient between any two species, and predator-prey theory (Holt Chapter 7) often assumes a constant attack coefficient for predators on prey. Or, if theory allows for changes in coefficients, often it

does not explicitly account for how adaptive behavior might influence key coefficients. Of course, there are exceptions; a subfield of theoretical ecology focuses explicitly on effects of adaptive behavior on population and community dynamics (Abrams 2000; Holt and Kimbrell 2007; Kotler and Brown 2007).

Proposition 6

Foraging theory requires foragers to estimate parameters that influence the fitness associated with different foraging options. Simple foraging models assume, in essence, that foragers are omniscient – that they have precise, complete information about the parameters needed to apply foraging theory. In reality, of course, foragers generally have only imprecise, incomplete estimates of the relevant parameters (Stephens and Krebs 1986; Sih 1992). Thus foragers must either make decisions based on incomplete information, or they can sample, learn and enhance the quality of their estimates of relevant parameters. Incomplete information and uncertainty are general problems in evolutionary ecology that shape all aspects of behavior and life histories (Dall et al. 2005), as well as physiology and morphology. In a later section, I summarize two main ways that foraging theory has addressed forager uncertainty: 1) sampling to get better information; and 2) risk sensitivity (here, risk = uncertainty) to either prefer or avoid options with more uncertain outcomes.

AN OVERVIEW OF FORAGING MODELS

Here, I provide a brief overview on major types of models in foraging theory from early simple models to more complex ones that include many aspects of reality. En route, I identify several major predictions of foraging theory that are listed as propositions about outcomes.

The basic optimal patch use and diet models

Modern foraging theory began in the 1960s with a set of models by, in particular, Robert MacArthur and colleagues on optimal diets and optimal patch use (MacArthur and Pianka 1966; Emlen 1966). These models were important in formalizing the notion of using the optimality approach to explain foraging behavior. In the 1970's, a new set of optimal foraging models generated the intuitively reasonable predictions that have become part of standard ecological dogma (Schoener 1971; Werner and Hall 1974; Charnov 1976a,b). In 1986, Stephens and Krebs

published a seminal book that summarized developments in foraging theory up to that point, and in 2007, an edited volume on foraging behavior and ecology brought the state of the field up to date (Stephens et al. 2007).

85

The basic optimal patch use model pictures a forager moving through a series of preycontaining patches separated by areas without prey. Following the basic propositions listed earlier: 1) optimal patch use theory proposes that foragers attempt to maximize net energy intake; 2) the forager's focal behavior is T = the time spent foraging in each patch; 3) the forager's mean net energy intake is g(T)/(T + t) where g(T) is its cumulative net energy intake as a function of T, and t is the mean transit time between patches. Either because foragers deplete prey or because prey go into hiding, the forager's instantaneous net energy intake rate while in the patch (dg/dT) goes down with time spent in the patch (Charnov et al. 1976). In economic terminology, the *marginal value* of the patch decreases over time. At some point, the marginal value of the patch is low enough that the forager should leave; however, when a forager leaves a patch, it must go through a 'transit time' without food before entering the next patch. That is, the forager faces a tradeoff. The next patch might have a higher initial net energy intake rate; however, to get to the next patch requires a 'down time' with no energy intake.

This problem is an example of a general 'stay versus leave' paradigm where a forager is in a current patch and must decide, at each moment, whether to stay or move on to the next patch. In principle, the decision is simple. The forager should compare its current net energy intake rate (its reward rate) to its expected average rate if it left the patch. Its optimal patch use should then fit the following prediction.

Proposition 7

A forager should stay in a patch as long as its current reward rate (the 'marginal value' of the patch) is higher than the average reward rate for the rest of the habitat, and it should leave as soon as the current reward rate is no better than the expected rate elsewhere. Charnov (1976b) called this the 'marginal value theorem'. It is one of the best know theoretical ideas in ecology. See Figure 4.2 for a graphical depiction of this theorem.

Corollary predictions are as follows. 1) In a given habitat, foragers should stay longer in patches that initially had higher prey availability. This prediction is obvious and did not need theory to say so. 2) In a given habitat, the marginal value of all patches (i.e., prey abundance in

all patches) should be the same at the point when foragers leave the patch. If we assume that foragers do not know which patch is coming next, then in all patches, foragers should stay as long as the current patch's reward rate is better than the average for the overall habitat; i.e., all patches use the same criterion for deciding when to leave a patch. After foragers have completed a bout, all patches should be reduced to the same food density. Or, as Brown and Kotler (2007) have emphasized, all patches should have the same 'giving up density (GUD)' – the same prey density at which foragers give up and leave that patch. Conversely, the observed GUD in all patches should reflect the quality of the overall habitat (GUD should be lower in lower quality habitats) and can thus be used as an indirect indicator of the forager's assessment of habitat with longer transit times (e.g., if patches are further apart) should spend more time in patches of a given quality (than foragers in habitats with shorter transit times). Literature reviews suggest that in situations where reality approximates the patch use model's assumptions, optimal patch use theory does a good job of explaining observed patch use (Stephens and Krebs 1986; Nonacs 2001).

Note that the basic 'stay versus leave' paradigm and the predictions of optimal patch use theory can, in principle, apply to any economic decision-making involving the choice among options that feature diminishing returns. For example, in our own lives, this theory and its predictions could apply for time spent persisting in a particular job, project, or partnership.

The basic optimal diet model (Schoener 1971; Werner and Hall 1974; Charnov 1976a) also, in essence, features a 'stay versus leave' framework. When a particular prey item is encountered, the forager must decide whether to accept/attack it (stay with it) or reject/ignore it (leave and continue searching for other prey). Each prey type has a net energy gain associated with attacking it (which depends on its energy content, the energy cost of capturing and consuming it, and the probability of capture), and a handling time (a period of time spent on that prey item when the forager cannot search for and capture other prey). The forager should attack a prey item if its net energy intake per unit handling time is greater than the expected net energy intake if it continued searching. Put another way, when a forager attacks and handles a prey item, it suffers an opportunity cost in the sense that it loses an opportunity to possibly encounter and attack other prey. The forager should attack a prey item if the benefit of that item is larger than the opportunity cost. If we rank prey by prey value, the prey type with highest value should

always be taken when encountered. Since no better options exist, there is nothing to be gained from rejecting the best option to wait for something better to appear. In contrast, a lower value prey item should be rejected if higher value prey are abundant enough; i.e., if there is a high enough chance that during the time when the forager is handling the low value prey, a high value prey item might appear. The general prediction is as follows.

Proposition 8

When high value prey are abundant, foragers should specialize on high value prey (i.e., reject low value prey), whereas when high value prey are scarce, foragers should have broad, generalized diets (i.e., they should attack both low and high value prey when they are encountered). Again, this basic paradigm and predictions should also apply for other choices, e.g., mate choice (Crowley et al. 1991).

A literature review by Sih and Christensen (2001) found that optimal diet theory's predictions usually worked well for foragers feeding on prey that are relatively immobile (e.g., flowers, fruits or relatively immobile animals), but that the theory performed poorly for foragers on mobile prey (e.g., large fish on small fish, large insects on small insects, mammalian predators on mammalian prey). The authors suggested that this is because with mobile prey, prey behavior (and not predator active diet choice) plays a major role in determining predator diets. With mobile prey, many prey are hiding and/or difficult to capture, whereas only some prey are both exposed and easy to capture. When most prey are hiding and difficult to capture, predators often attack any prey that they encounter that could conceivably be captured. The diets for predators on mobile prey then tend to be dominated not necessarily by prey with high energy per unit handling time, but by prey that are exposed and relatively easy to capture. To emphasize, this observation does not contradict basic optimal diet theory. Instead, it suggests that to understand the optimal diets of predators on mobile prey, empiricists must account for variation in prey exposure to predators, and the effect of prey escape ability in determining prey value.

Adding complexities and realities

From the 1970's on, foraging theory acknowledged and incorporated many aspects of reality that were not included in the early models. That is, as a field, foraging theory matured by adding

new subfields via the classic theory-empirical test feedback loop. Empirical work pointed out key missing complexities that were then incorporated into more sophisticated models which led to new tests that revealed new complexities and so on. In terms of the principles of ecology, basic foraging theory was built on applying principle 8 (evolutionary causes) to principles 1, 5 and 6 (heterogeneous distributions of organisms and their finite resources, as well as of the environment, in general). As the field developed, it added theory and tests associated with principles 4 (contingency and the need to account for uncertainty) and 7 (biotic interactions including competition and predation risk). In the next sections, I review some of these main developments, and in Future Directions, I note some connections to other areas of ecology that remain largely unexplored.

Lack of information

Early models assumed that foragers are omniscient – that they have complete and precise information about their environment – specifically, that they have accurate estimates of all parameters needed to apply optimal foraging theory. Obviously, this is presumably rarely if ever true. Instead, foragers must both estimate the relevant benefit-cost parameters (Proposition 6) and given uncertainty in parameter estimates, make decisions that account for the uncertainty. The challenge for theory has been to classify several main types of responses to lack of information and to capture their essence in simple, testable models.

One general type of theory on lack of information revolves around foragers sampling and learning in order to forage more efficiently in the future. Optimal sampling balances the benefits and costs of sampling. The benefit lies in the value of having more accurate parameter estimates that should result in higher foraging efficiency in the future. This benefit can depend on the interplay between environmental factors (e.g., inherent predictability and stability of the environment, one cannot learn much useful about environments that are highly unpredictable and unstable) and the organism's sensory and cognitive abilities. The cost of sampling can include temporarily reduced foraging efficiency while sampling items or patches that an omniscient forager would not use, and risks taken that an omniscient forager would not take (Sih 1992). Statistical decision theory (using Bayes' theorem) where foragers have a prior estimate of the world that is continually updated as new information comes in has been a useful framework for

analyses of optimal sampling (DeGroot 1970; McNamara 1982; Stephens and Krebs 1986; Sih 1992).

Optimal sampling regimes have been analyzed in a range of scenarios including the following. 1) Key parameters (e.g., prey value) are fixed, but foragers need to sample items to ascertain which are good versus poor items. Signal detection theory can be used to analyze effects of variation in the ease of discriminating the relative quality of different options (Getty et al. 1987; Stephens 2007). 2) Patches exhibit diminishing returns, but foragers need to assess the rate of diminishing returns and the patches' current marginal value (Oaten 1977; McNamara 1982). And, 3) options (patches or prey types) can be either good or bad, but these can change over time. The forager must sample to assess each option's current value (Stephens and Krebs 1986). Most models have examined changes in the prey regime, but some have looked at how uncertainty about predation risk influences foraging behavior (Sih 1992; Bouskila and Blumstein 1992; Abrams 1994). Dall et al. (2005) provide a general overview on information and its use in evolutionary ecology. As with other aspects of foraging theory, optimal sampling can also be applied to human decision making (e.g., via adaptive management). We should all plan our lives to not only maximize current net gains, but also to gather the information required to make more intelligent decisions in the future.

An alternative response to uncertainty is to incorporate risk, in the sense of uncertainty, into the evaluation of an options' value (Real and Caraco 1986; Stephens and Krebs 1986; Houston and McNamara 1999). Picture, for example, a world with two options A and B, where A yields a guaranteed, moderate reward rate, and B yields the same average reward rate as A, but with high variance. B is a 'boom or bust' option that could either yield a much higher or much lower reward rate. Drawing directly from economic utility theory (e.g., Keeney and Raiffa 1993), animals should prefer the uncertain, risky option B (they should be 'risk prone') if the increased benefit of the higher reward rate is greater than the cost of the lower reward rate. Conversely, foragers should prefer the low variance option A (they should be 'risk averse') if the increased cost of the lower reward rate outweighs the benefit of the higher reward rate. In nature, the foraging rate versus fitness function is probably often sigmoid-shaped where a range of low feeding rates yield little or no fitness, and high feeding rates yield diminishing returns in terms of fitness (and for humans, very high feeding rates result in obesity and reduced fitness).

Proposition 9

With a sigmoid fitness function, foragers with low average feeding rates have little to lose, and should thus be risk prone while foragers with high average feeding rates have little to further gain, and should thus be risk averse. In principle, these ideas could also explain risk sensitivity and gambling or investment strategies in humans.

Experimental tests of risk sensitivity theory have yielded mixed results (Kacelnik and Bateson 1997). Interestingly, foragers appear to treat uncertainty about time versus reward size differently. They tend to be risk averse about reward size, but risk prone about delay to reward (Kacelnik and Bateson 1997). Being risk prone about delays means that foragers, including many humans, tend to be impulsive – we strongly prefer immediate rewards to a degree that is difficult to explain using standard optimality theory (Stephens 2002). Extreme impulsiveness is associated with poor performance in school, and addiction (Mischel et al. 1989). Explaining impulsiveness and other aspects of risk sensitivity suggests a need to incorporate more about cognitive and perhaps neuroendocrine aspects of behavior into foraging theory.

State-dependent dynamic optimization

Interestingly, although we know that animals, including ourselves alter their foraging behavior depending on hunger level, or energy reserves, basic optimal foraging theory does not address this point. That is, early optimal foraging theory was not state-dependent. Seminal papers and books by McNamara and Houston (1986) and Mangel and Clark (1988) brought state-dependent dynamic models into the mainstream of foraging theory. These models address a particular key aspect of Proposition 4 – that foraging behavior depends on the forager's state.

In the foraging context, a state variable is a property of an individual that carries over across time, with a future value that is affected by current behavior. A dynamic feedback occurs where current state affects the optimal current behavior which affects future state and so on. Examples of ecologically-relevant state variables include energy reserves, size, reproductive value or other forms of assets, condition, vigor or information state. This framework adds several major aspects of reality into foraging theory: 1) as noted, although individual state obviously affects behavior, this was not included in earlier models; 2) state-dependence connects current short-term behavior to long-term fitness which is critical for putting risk and foraging needs into a common currency (a key issue for analyzing tradeoffs, Proposition 3); and 3) unlike

previous models, dynamic state-dependent models explicitly address temporal patterns of behavior – typically by building in a time horizon (e.g., an end-of-the-season) and predicting changes in behavior over time.

State-dependent models help to solve what had been termed the 'common currency' problem (McNamara and Houston 1986). Many empirical studies show that predation risk (or other sources of mortality) and energy needs both influence fitness and behavior. The challenge has been to clarify how foragers (and modelers) should balance risk against energy to identify the optimal behavior when these two are in different currencies: survival versus energy intake. How much energy should an animal give up for a unit of increased safety? This dilemma can be solved by converting energy gain into fitness terms. The problem is that the energy gained from short-term foraging often has little or no immediate effect on fitness. The fact that I skipped lunch to work on this paper had no discernible effect on my fitness. Thus a corollary problem is the need to have theory that connects short-term behavior to long-term energy budgets that affect long-term fitness. The solution is to account for state variables. Current foraging decisions affect energy reserves. In the short-term, energy reserves can immediately affect fitness if animals are close to starving. Over the long-term, the accumulation of many, small, short-term foraging decisions affects cumulative energy reserves that affect growth and future reproduction. To date, most state-dependent models feature relatively simplistic depictions of the relevant energetics; however, some include more sophisticated, realistic models of the biology of energy reserves (e.g. Brodin and Clark 2007).

A general result to emerge from dynamic optimization models is the 'asset protection principle' that says in essence that organisms should be more cautious when they have more to lose (Clark 1994).

Proposition 10

Animals that have low assets (e.g., low energy reserves, or low reproductive value, and thus low prospects for the future) should be bold (accept greater predation risk) to gain more energy, whereas animals with high assets should be cautious to protect their assets from catastrophic loss (e.g. death). Interestingly, over time, this negative feedback process should tend to reduce variation in assets and thus in asset-dependent behavior. Individuals that initially had lower assets should be more bold and active than those with initially more assets; thus individuals that

started with less should tend to catch up in assets to those who started with more (as long as they are not killed in the process).

Simple, basic foraging theory was agnostic about how foraging patterns might change over time. Should animals be more or less bold as they approach the end of a growing season, or the end of their lifetime? Another improvement offered by dynamic optimization theory is the fact that it generates predictions on how foraging behavior ought to change over the course of a day, a season or a lifetime. Opposite predictions emerge depending on whether the time horizon represents a lethal endpoint (e.g., onset of winter for an annual organism, or pond drying for a pond organism) or an opportunity to cash in assets (e.g., the onset of reproduction). General predictions are that: 1) if the goal is to survive until the organism can cash in assets, then individuals should take fewer risks (i.e., be more cautious to protect assets) as they approach that horizon; whereas 2) if the time horizon is a lethal endpoint, then individuals should take more risks as that endpoint approaches.

Accounting for state variables also reminds us that energy is not the only benefit to be gained from food. Food also contains a myriad of other nutrients (protein, minerals, vitamins) as well as toxins. Simple foraging theory assumed, in essence, that food contains no toxins, and that all beneficial aspects of food are positively correlated; i.e., that foods that yield more energy/handling time also yield more other nutrients/handling time. A simple approach that incorporates other nutrient considerations involves using linear programming which identifies the behavior that maximizes a benefit (or minimizes a cost) under the constraint that the optimal behavior must also satisfy other minimum needs (e.g., a minimum sodium need; Belovsky 1978). In reality, effects of the mix of nutrients on forager performance can be complex (Newman 2007). In principle, multiple nutrient considerations could be incorporated into a model by including multiple, fitness-related state variables. Another area of ecology that examines multiple nutrients is ecological stoichiometry (see Burke and Lauenroth Chapter 11). An exciting future step could thus involve blending principles of ecological stoichiometry with modern foraging theory.

Note that while dynamic, state-dependent models provide the significant advantage of adding many, important aspects of reality to earlier, simpler foraging models, this additional reality comes with costs. Because the models are more complex, and more specific, they are also less general. In addition, it is often difficult to get realistic estimates of the many parameters and

functions that are part of these models. In principle, one should do sensitivity analyses to examine effects of uncertain parameter estimates on predictions; however, often, the overall parameter space is so large that this is either difficult to do, or difficult to publish. Finally, with complex models with many interacting factors, it can be difficult to identify the biological mechanisms (as opposed to convenient mathematical assumptions) underlying a predicted pattern. Progress using dynamic, state-dependent models will likely depend on the balance between these benefits and costs.

Social and species interactions: competition, predation risk

Basic optimal foraging theory assumed, in essence, that predators forage alone. In reality, following fundamental principle 7 (birth and death rates depend on abiotic and biotic factors), foragers often, perhaps usually, must account for competitors, or predation risk (or other sources of mortality) or both. Accounting for interactions among competing foragers, and for predators or other species adds two major types of complexity to foraging models: 1) it makes the optimization criterion more complex (as discussed above), and 2) it brings in the notion of games where the best behavior for a forager depends on decisions made by others (competitors or predators) and vice versa (Maynard Smith 1982).

Beginning almost four decades ago, the field accounted for intraspecific competition via analyses of the 'ideal free distribution' (Fretwell and Lucas 1970). Picture a simple scenario with only two patches, A and B, where A initially has more resources than B. A single forager in a non-competitive world should prefer patch A; however, if too many competitors join that forager in patch A, competition might be so intense that A is no longer better than B.

Proposition 11

If foragers are free to move whenever they could do better elsewhere, then interestingly, at the ideal (optimal) free distribution, all patches should yield equal reward rates. The simple underlying logic is that if any patch is yielding higher reward rates than others, some competitors should move to that patch (thus increasing competition in that patch) until it is no better than other patches. Foragers should continue to move among patches until all patches yield equal reward rates; at that point, there is no benefit to be gained from moving. In evolutionary game theory terms, this is an evolutionarily stable strategy (ESS; Maynard Smith 1982). In the

simplest models, at the ideal free distribution (at the ESS), the ratio of consumers in two patches should 'match' the ratio of resource inputs in the two patches; e.g., if patch A has a four times higher resource input rate than patch B, then at the ESS, four times more competitors should be in patch A than patch B. This basic framework and predictions could, in theory, apply to the distribution of competitors among any set of options – including human choice of places to live, or jobs. For example, in theory, if humans followed the ideal free distribution, the quality of life should be equal in all cities. If the quality of life, including costs of living, was higher in San Francisco than Oklahoma City, people should move to San Francisco until it is so crowded that it is no longer better than Oklahoma City.

Many studies have confirmed that, in broad agreement with ideal free distribution theory, competitors are often more abundant in areas with more resources. While some of these studies have found that foragers indeed match their resources, most studies found 'undermatching' where the consumer ratio was less than the resource ratio (e.g., if the ratio of resource inputs in two patches was 4:1, the consumer ratio was only 2:1). Subsequent theory and experiments suggested that undermatching could be explained by: 1) aggressive interactions among competitors where dominant consumers keep less dominant individuals out of patches with more resources; 2) movement costs; and 3) imperfect knowledge (e.g., Abrahams 1986). The logic on the latter two points is that consumers might stay in patches that yield lower reward rates either because it is too costly to move to a better patch, or because they do not realize that other patches are better than their current patch.

Competition has also driven the evolution of alternative competitive foraging strategies within each patch (Giraldeau and Caraco 2000; Waite and Field 2007). For example, in some systems, subordinate individuals serve as 'producers' who search for food, while dominant individuals are 'scroungers' who steal food first found by producers. In essence, scroungers parasitize the efforts of producers. This scenario is also often seen and referred to as 'piracy' across species.

Also beginning several decades ago, both modeling and empirical work incorporated predation risk into analyses of foraging behavior (Sih 1980; Werner et al. 1983).

Proposition 12 $\$

Foraging behavior often balances the conflicting demands of gaining energy and avoiding predation risk. Options (patches, diet items) that yield higher energetic rewards are often also dangerous. Risk then affects foraging decisions, and feeding (energy) demands affect predator avoidance (Lima and Dill 1990: Lima 1998; Brown and Kotler 2007; Bednekoff 2007). As noted earlier, state-dependent dynamic optimization models provide insights on how foragers might balance these tradeoffs. Other notable realities about foraging /risk tradeoffs that have been modeled and studied empirically include: 1) forager responses to multiple predators, particularly those that require different antipredator responses (Lima 1992; Matsuda et al. 1994; Sih et al. 1998); 2) effects of temporally varying predation risk (Lima and Bednekoff 1998; Ferrari et al. 2009); 3) ideal free distributions with predation risk (McNamara et al. 1999; Grand et al. 1999); and 4) predator-prey space use when predator attempts to use areas with more prey can be offset by prey attempts to avoid predators (Heithaus 2001; Lima 2002; Alonzo 2002; Hammond et al. 2007).

Overall, in the four decades since optimality modeling began to be applied regularly in ecology, foraging theory has grown in a steady, organic fashion, where in each stage, existing theory has been criticized for being too simple, leading to new theory that incorporated important aspects of reality. Though some would dispute this assessment, my sense is that the resulting body of theory has been reasonably successful at explaining, or even predicting major patterns in foraging behavior (Stephens et al. 2007).

LINKING FORAGING THEORY TO POPULATION AND COMMUNITY ECOLOGY

Foraging theory provides a critical mechanistic bridge between organismal ecology and several other fields of ecology (Figure 4.3). Most theory on population and community dynamics assumes that species interactions are important. The persistence and coexistence of species are thought to depend, at least in part, on competition and predator-prey interactions (MacArthur 1972; Chesson 2000; Kotler and Brown 2007). One research goal is thus to quantify the strength of these interactions and their impacts on ecological patterns. For a mechanistic ecologist, a key additional step is to explain or even predict the strength of these interactions in terms of species traits. The strength of competition (as measured by a competition coefficient) depends on species niches (diets, habitat use; Chase, this book) that, in turn, depend on species traits (e.g.,

morphology, physiology, life history and behavior), and the strength of the predator-prey interaction (e.g. the predator attack coefficient) depends on predator and prey traits (Murdoch and Oaten 1976). Foraging theory provides an explicit, quantitative basis for bridging from species traits to the interaction strengths that underlie population and community dynamics. Notably, while most simple population/ community models assume that key interaction rates are constants, in reality, they invariably change with prey, competitor or predator density in ways that depend on behavior. A central goal of foraging theory has been to explain or even predict these changes.

Proposition 13

Foraging theory provides a mechanistic basis for understanding niches and patterns of competition that underlie patterns of coexistence. Indeed, foraging theory was first created by MacArthur and colleagues (e.g., MacArthur and Pianka 1966) explicitly to address this proposition quantitatively.

A flagship example of the use of foraging theory to explain competitive coexistence is the classic work of Werner and colleagues on diets, habitat use and competition among North American sunfish. They used optimal diet theory to predict the diets of sunfish in the laboratory (Werner and Hall 1974) and field (Mittelbach 1981). Because sunfish species differ in body shape and jaw morphology, they differ in efficiency on different food types. These differences can then explain their feeding niches and how they shift in the presence of competitors (Werner 1977). Smaller sunfish, however, suffer predation risk from predatory bass. Extending theory to include size-dependent foraging-risk tradeoffs (Werner and Gilliam 1984) predicted size-dependent changes in habitat use (ontogenetic niche shifts) from safer, vegetation (thus feeding on vegetation prey) to open water (feeding on zooplankton). These predictions were corroborated with field experiments that manipulated the presence/absence of bass (Werner et al. 1983).

Attempts to combine foraging theory and competition ran into an interesting paradox. According to foraging theory, when competition reduces food availability, foragers should broaden their diets, and become generalists. Competitors should thus increase in niche overlap. In contrast, competition theory (the 'compression hypothesis') suggested that when resources are scarce, species should reduce niche overlap with competitors by specializing on preferred

resources. In recent years, renewed interest in the connection between foraging theory and competition has arisen around new attempts to explain this paradox (Robinson and Wilson 1998; Svanback and Bolnick 2005, 2007). One possible resolution to this paradox hypothesized that when resources become scarce, foragers should avoid competition by becoming habitat specialists while remaining a resource generalist within their preferred habitat. Another solution suggests the possibility that all competitors prefer the same high value prey types when they are abundant, but that under strong competition, competitors broaden their diets, but diverge to include different less preferred prey types as shared, high value prey become scarce.

Along similar lines, foraging theory has been used to predict adaptive shifts in predator and prey behaviors that influence predator-prey dynamics (Gleeson and Wilson 1986; Abrams 2000; Holt and Kimbrell 2007; Brown and Kotler 2007). An exciting, recent offshoot of this work emphasizes the ecological importance of trait-mediated, non-consumptive effects of predators on prey (Preisser et al. 2005; Schmitz et al. 2008; Orrock et al. 2008). That is, even if predators rarely consume prey, predators can have major impacts on prey fitness and prey intraspecific and interspecific interactions by causing shifts in prey habitat use, activity and diets (Sih 1990; Lima 1998; Brown and Kotler 2007). These non-consumptive effects can cascade down to influence the prey's prey. When carnivores cause shifts in herbivore behavior, this can have large impacts on plants (Power and Matthews 1985; Schmitz 2004).

Proposition 14

Predators often have strong non-consumptive effects on communities that are mediated by shifts in foraging behavior. Strong evidence in support of this proposition comes from a meta-analysis that found that non-consumptive effects of predators on prey (or on prey of prey) are often stronger than direct, consumptive effects of predation (Preisser et al. 2005). Foraging theory (via the risk-foraging tradeoff) can predict the nature and magnitude of trait-mediated, non-consumptive effects (Abrams 2000; Brown and Kotler 2007).

Another offshoot of foraging theory that can potentially interface with population and community ecology involves using foraging theory to predict optimal movement or dispersal among patches. In a world where many, perhaps most habitats are patchily distributed (fragmented), a key issue for understanding species persistence and community structure is the interplay between interactions within patches and dispersal between patches. In the 1990's, a

major area of study for both basic and applied ecology (e.g., conservation ecology) was metapopulation ecology (Hanksi 1999), population ecology in a patchy world where local extinctions can be offset by recolonizations. In the 2000's, this field was broadened to include multi-species interactions; i.e., metacommunity ecology (Holyoak et al. 2005; Leibold, this book). Theory in both metapopulation and metacommunity ecology often assumes that dispersal rates are fixed. In fact, in many animals, both dispersal from one patch and settlement in another almost certainly depend on habitat quality in both patches, as well as in the matrix habitat between patches. In turn, habitat quality depends on multiple factors including resource levels, competition and predation risk. Foraging theory that predicts predator, competitor and prey behavior as a function of these ecological factors should thus play a useful role in predicting patterns of adaptive dispersal.

Overall, the key insight from foraging theory (and more generally from evolutionary ecology) for the rest of ecology is the simple notion that behavioral plasticity (often adaptive behavior) is omnipresent and it matters. Behavior generates non-linearities (due to dynamic changes in rates) that can affect dynamics and outcomes in ways that cannot be predicted without accounting for the plasticity. The general suggestion is thus to use foraging theory to provide a mechanistic, adaptive understanding of how plasticity in response to tradeoffs underlies ecological patterns. Behavioral plasticity is a fundamental aspect of the 'theory of organisms' (Zamer and Scheiner, in prep); ignoring behavioral plasticity will often prove to be unwise.

FUTURE DIRECTIONS

Although foraging theory logically provides a mechanistic basis for explaining behavioral responses to resources and/or predators that are ubiquitous and potentially important – often, this behavioral plasticity is not accounted for explicitly in other ecological fields. Why not? The answer probably lies in the fact that it is an additional complexity that can be hard to incorporate in an already full research program. Ecologists often have a hard enough time getting sufficient data on their preferred level of study (i.e., the population, community or ecosystem level) without having to also study and understand phenomena at the level of individual behavior. And yet, to me, it is difficult to argue against the notion that consumers exhibit behavioral plasticity and that that plasticity is important. A 'solution' to this practical problem is to gloss over the details on foraging behavior, but to still incorporate the main patterns. While we know that

foragers often show adaptive responses, theory suggests that depending on model details, those responses and their impacts could go in diametrically opposite directions (Abrams 2000 and numerous other papers by Abrams and others). Additional modeling and empirical work is still needed to clarify key *generalities* about flexible foraging behaviors, in particular, those that appear to be ecological important.

Interestingly, new empirical methods hold great promise for aiding the endeavor of incorporating behavioral plasticity into ecology. To date, many of the most detailed foraging studies have been done in the laboratory where it is relatively easy to observe animal behavior. Until recently, for many mobile animals it has been difficult, if not impossible, to get good field data on foraging behavior. New technologies (e.g., GPS collars, multi-camera video systems with automated data analyses, other distributed remote sensing systems), however, are allowing us to remotely monitor more individuals than ever, with more information collected per individual, over longer periods of time, along with simultaneous monitoring of key environmental factors.

In addition, new statistical methods should help generate more accurate and nuanced insights on foraging behavior. To date, most patch use experiments have adopted the simplistic, model-derived scenario of a set of equal-sized patches of food separated by areas with no food. Reality often features much more complex landscapes. New spatial statistics and landscape methods can be used to more rigorously characterize realistic patterns of spatial and temporal heterogeneity in resources, foragers and predators, and to more rigorously identify patterns of response. Also, to date, most foraging studies have used standard p-value driven statistical methods to compare the predictions of one or a few foraging models against null expectations. An alternative approach uses model choice methods (Burnham and Anderson 2002) to compare numerous alternative models in order to identify behavioral rules that best explain observed predator and prey behavioral patterns (e.g., Luttbeg and Langen 2004; Hammond et al. 2007).

Another exciting future direction involves linking foraging theory to conceptual areas in ecology that are not currently well linked to foraging theory. These can be identified by looking at some of the other chapters in this book. Ecological stoichiometry focuses on how elemental mismatches (e.g., involving carbon: nitrogen: phosphorus ratios) between consumers and resources affect their ecology (Sterner and Elser 2002; Burke and Lauenroth Chapter 11). Although most of foraging theory uses energy (carbon) as the key measure of food value,

nutrient constraints are clearly important and have been included in foraging theory (Newman 2007). Principles of ecological stoichiometry could provide a mechanistic basis for predicting how prey with different elemental ratios should differ in prey value depending on the consumer's stoichiometry. The notion that multiple chemical constituents in food should matter can be incorporated into foraging theory using dynamic, state-dependent models with multiple elements included as multiple state variables. Foraging theory could then help to clarify how prey elemental content and stoichiometry (along with handling times) might balance with fitness benefits of different elements and risks of mortality to predict foraging behavior.

The metabolic theory of ecology (Brown et al. 2004) posits that basic constraints of size, temperature, and the need to move resources within organisms in vascular networks with constrained geometries might affect not just metabolic rates, but also numerous aspects of the organisms' general ecology and evolution. The basic assumption is that size and temperature play the key roles in driving metabolic rates and other outcomes. Standard foraging theory does not dispute the notion that foraging behavior likely depends on predator (and prey) size and on temperature; however, it does not typically draw on first principles to predict how size and temperature should influence foraging behavior. Thus foraging theory could gain from metabolic theory. Conversely, most of the metabolic theory's models assume that resource limitation per se and foraging behaviors that govern resource intake rates do not limit metabolic rates. While metabolic theory might explain some of the major patterns in the natural world, much of the residual variation might be explained by patterns of foraging and resource limitation.

Ecosystem ecology, in general, examines fluxes of energy and nutrients (water, carbon, nitrogen, phosphorus, various elements) among major abiotic and biotic categories (e.g., among trophic levels). Ultimately, when fluxes involve consumers and resources, rates of flux should depend on individual foraging decisions. While there are some examples of where shifts in foraging behavior fundamentally change nutrient cycles or trophic dynamics (see Schmitz et al. 2008), most foraging ecologists ignore ecosystem processes and vice versa. Clearly, developing this bridge holds great promise.

Finally, foraging theory remains poised to contribute to various fields of applied predator-prey ecology, including farm animal foraging and meat production, fish foraging and fisheries, foraging by animal disease vectors, foraging by biological control agents, and foraging

by species of conservation concern. For each issue, foraging theory can provide a better understanding of predator and/or prey behavior that can, in turn, potentially help enhance the efficacy of management.

A particularly exciting area of future study in conservation ecology and invasion ecology involves understanding variation among foragers (predators or prey) in their response to novel environments, which in the modern world often consists of human-induced rapid environmental changes. In the predator-prey context, a key issue involves prey responses to novel predators and predator responses to novel prey. Invasive, exotic predators often have major negative impacts on prey, but in other cases, prey cope well with exotic predators (Cox and Lima 2006). (Mack et al. 2000; Cox and Lima 2006; Salo et al. 2007). An important question is - why do some prey respond well to exotic predators while others do not (e.g., Rehage et al. 2005)? Conversely, some foragers readily utilize novel resources (e.g., some herbivores utilize crops and ornamental plants) while others do not. What explains the variation in response to novel foods (e.g., Rehage et al. 2005)? Interestingly, addressing the issue of responses to exotic, invasive species will require a twist on the usual foraging theory. Foraging theory normally attempts to explain behaviors that are adaptive, presumably due to a long evolutionary history of dealing with similar situations. Now, our question is - why do some individuals respond well and others respond poorly to predators or prey in situations that they have never seen before? Traditional foraging theory might still play a useful role in explaining how responses to cues that in the past provided good indicators of prey value or predation risk might now sometimes result in mismatches that lead predators or prey to exhibit inappropriate behaviors. These mismatches have been termed ecological traps (Schlaepfer et al. 2004). In a confusing modern world, an important task for foraging theory might be to provide a jumping off point for understanding such traps.

Overall, after more than four decades, foraging theory remains a vital area of study with many exciting opportunities for new bridges with other constitutive theories in ecology. Insights from foraging theory on behavioral plasticity remain to be fully integrated into other areas, and insights from other areas continue to hold promise for further refining of foraging theory.

Literature Cited

- Abrahams, M.V. 1986. Patch choice under perceptual constraints: A case for departures from an ideal free distribution. Behavioral Ecology and Sociobiology 19:409-415.
- Abrams, P.A. 1994. Should prey overestimate the risk of predation? American Naturalist 144:317-328.
- Abrams, P.A. 2000. The evolution of predator-prey interactions: Theory and evidence. Annual Review of Ecology and Systematics 31:79-105.
- Alonzo, S.H. 2002. State-dependent habitat selection games between predators and prey: the importance of behavioural interactions and expected lifetime reproductive success. Evolutionary Ecology Research 4: 759-778.
- Bednekoff, P.A. 2007. Foraging in the face of danger. Pages 305-330 in D.W. Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Belovsky, G.E. 1978. Diet optimization in a generalist herbivore: The moose. Theoretical Population Biology 14:105-134.
- Bouskila, A. and D.T. Blumstein. 1992. Rules of thumb for predation hazard assessment: predictions from a dynamic model. American Naturalist 139:161-176.
- Brodin, A. and C.W. Clark. 2007. Energy storage and expenditure. Pages 221-269 in D.W. Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Brown, J.H., J.F. Gillooly, A.P. Allen, V.M. Savage, G.B. West. 2004. Towards a metabolic theory of ecology. Ecology 85:1771–1789.
- Brown, J.S. and B.P. Kotler. 2007. Foraging and the ecology of fear. Pages 437-480 in D.W.Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Burnham, K.P. and D.R. Anderson 2002. Model Selection and Multi-model Inference. Springer-Verlag Press.
- Charnov, E.L. 1976a. Optimal foraging: attack strategy of a mantid. American Naturalist 110:141-151.
- Charnov, E.L. 1976b. Optimal foraging: the marginal value theorem. Theoretical Population Biology 9:129-136.

- Charnov, E.L., G.H. Orians and K. Hyatt. 1976. Ecological implications of resource depression. American Naturalist 110:247-259.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31:343-366.
- Clark, C.W. 1994. Antipredator behavior and the asset protection principle. Behavioral Ecology 5:159-170.
- Clark, C.W. and M. Mangel 2000. Dynamic State Variable Models in Ecology: Methods and Applications. Oxford University Press, Oxford.
- Cox, J. G., and S. L. Lima. 2006. Naivete and an aquatic-terrestrial dichotomy in the effects of introduced predators. Trends in Ecology & Evolution **21**:674-680.
- Crowley, P.H., S.E. Travers, M.C. Linton, S.L. Cohen, A. Sih, and R.C. Sargent. 1991. Mate density, predation risk and the seasonal sequence of mate choices: a dynamic game. American Naturalist 137:567-596.
- Dall, S.R.X., L.A. Giraldeau, O. Olsson, J.M. McNamara and D.W. Stephens. 2005.
 Information and its use in evolutionary ecology. Trends in Ecology and Evolution 20:187-193.
- DeGroot, M.H. 1970. Optimal Statistical Decisions. McGraw-Hill, New York.
- DeWitt, T.J. and SM Scheiner. 2004. Plasticity. Functional and Conceptual Approaches. Oxford University Press, Oxford.
- Emlen, J.M. 1966. The role of time and energy in food preference. American Naturalist 100:611-617.
- Ferrari, M.C.O., A. Sih and D.P. Chivers. 2009. The paradox of risk allocation: a review and prospectus. Animal Behaviour, in press.
- Fretwell, S.D. and H.L. Lucas. 1969. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheoretica 18:16-36.

Fryxell, J.M. 1998. Individual Behavior and Community Dynamics. Chapman & Hall, London.

- Getty, T., A.C. Kamil and P.G. Real. 1987. Signal detection theory and foraging for cryptic and mimetic prey. Pages 525-548 in A.C. Kamil, J.R. Krebs and H.R. Pulliam, eds. Foraging Behavior. Plenum Press, New York.
- Giraldeau, L.-A. and T. Caraco 2000. Social Foraging Theory. Princeton University Press, Princeton, NJ.

- Gleeson, S.K. and D.S. Wilson 1986. Optimal foraging and prey coexistence. Oikos 46:139-144.
- Grand, T.C. and L.M. Dill. 1999. Predation risk, unequal competitors and the ideal free distribution. Evolutionary Ecology Research 1:389-409.
- Hammond, JI, BT Luttbeg and A. Sih. 2007. Predator and prey space use: dragonflies and tadpoles in an interactive game. Ecology 88:1525-1535.

Hanski, I. 1999. Metapopulation Ecology. Oxford University Press, New York,

- Heithaus, M.R. 2001. Habitat selection by predators and prey in communities with asymmetrical intraguild predation. Oikos 92:542-554.
- Holt, R.D. and T. Kimbrell. 2007. Foraging and population dynamics. Pages 365-395 in D.W.Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Holyoak, M., M.A. Leibold and R.D. Holt. 2005. Metacommunities: Spatial Dynamics and Ecological Communities. University of Chicago Press, Chicago.
- Houston, A.I. and J.M. McNamara. 1999. Models of Adaptive Behaviour. Cambridge University Press, Cambridge.
- Hubbell, SP. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ.
- Hugie, D.M. and L.M. Dill 1994. Fish and game: a game theoretic approach to habitat selection by predators and prey. Journal of Fish Biology 45:151-169.
- Kacelnik, A. and M. Bateson 1997. Risk-sensitivity: crossroads for theories of decision making. Trends in Cognitive Sciences 1:304-309.
- Keeney, R.L. and H. Raiffa. 1993. Decisions with Multiple Objectives: Preference and Value Tradeoffs. Cambridge University Press, Cambridge.
- Kotler, B.P. and J.S. Brown. 2007. Community ecology. Pages 397-434 in D.W. Stephens, J.S. Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Lima, S.L. and L.M. Dill 1990. Behavioral decisions made under the risk of predation: a review and prospectus. Canadian Journal of Zoology 68:619-640.
- Lima, S. L. 1992. Life in a multipredator environment some considerations for antipredatory vigilance. Annales Zoologici Fennici 29:217-226.

Lima, S.L. 1998. Stress and decision making under the risk of predation: recent developments from behavioral, reproductive and ecological perspectives. Advances in the Study of Behavior 27:215-290.

- Lima, S.L. 2002. Putting predators back into behavioral predator-prey interactions. Trends in Ecology and Evolution 17:70-75.
- Lima, S.L. and P.A. Bednekoff 1999. Temporal variation in danger drives antipredator behavior: the predation risk allocation hypothesis. American Naturalist 153:649-659.
- Ludwig, D. and L. Rowe 1990. Life history strategies for energy gain and predator avoidance under time constraints. American Naturalist 135:686-707.
- Luttbeg, B.T. and A. Sih. 2004. Predator and prey habitat selection games: the effects of how prey balance foraging and predation risk. Israel Journal of Zoology 50: 233-254.
- Luttbeg B.T. and T.A. Langen. 2004. Comparing alternative models to empirical data: Cognitive models of western scrub-jay foraging behavior. American Naturalist 163: 263-276.
- MacArthur, R.H. 1972. Geographical Ecology. Harper and Row, New York.
- MacArthur, R.H. and E.R. Pianka 1966. On optimal use of patchy environment. American Naturalist 100:603-609.
- Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. Ecological Applications 10:689-710.
- Mangel, M. and C.W. Clark 1988. Dynamic Modeling in Behavioral Ecology. Princeton University Press, Princeton, NJ.
- Matsuda, H., M. Hori, and P. A. Abrams. 1994. Effects of predator-specific defense on community complexity. Evolutionary Ecology 8:628-638.
- Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.
- McNamara, J.M. 1982. Optimal patch use in a stochastic environment. Theoretical Population Biology 21:269-288.
- McNamara, J.M. and A.I. Houston 1986. The common currency for behavioral decisions. American Naturalist 127:358-378.
- Mischel, W., Y. Shoda and M.L. Rodriguez. 1989. Delay of gratification in children. Science 244:933-938.

Mittelbach, G.G. 1981. Foraging efficiency and body size: a study of optimal diet and habitat use by bluegills. Ecology 62:1370-1386.

- Moody, A.L., A.I. Houston and J.M. McNamara. 1996. Ideal free distributions under predation risk. Behavioral Ecology and Sociobiology 38: 131-143.
- Murdoch, W.W. and A. Oaten 1975. Predation and population stability. Advances in Ecological Research 9:1-131.
- Newman, J. 2007. Herbivory. Pages 175-218 in D.W. Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Nonacs, P. 2001. State-dependent behavior and the marginal value theorem. Behavioral Ecology 12:71-83.
- Oaten, A. 1977. Optimal foraging in patches: a case for stochasticity. Theoretical Population Biology 12:263-285.
- Orrock, J.L., J.H. Grabowski, J.H. Pantel, S.D. Peacor, B.L. Peckarsky, A. Sih and E.E. Werner. 2008. Consumptive and nonconsumptive effects of predators on metacommunities of
- Pierce, G.J. and J.G. Ollason 1987. Eight reasons why optimal foraging theory is a complete waste of time. Oikos 49:111-118.
- Power, M. E., W. J. Matthews, and A. J. Stewart. 1985. Grazing minnows, piscivorous bass, and stream algae - Dynamics of a Strong Interaction. Ecology 66:1448-1456.
- Preisser, E. L., D. I. Bolnick, and M. F. Benard. 2005. Scared to death? The effects of intimidation and consumption in predator-prey interactions. Ecology 86:501-509.
- Pulliam, H.R. 1975. Diet optimization with nutrient constraints. American Naturalist 109: 765-768.
- Real, L. and T. Caraco. 1986. Risk and foraging in stochastic environments. Annual Review of Ecology and Systematics. 17: 371-390.
- Rehage, J. S., B. K. Barnett, and A. Sih. 2005. Behavioral responses to a novel predator and competitor of invasive mosquitofish and their non-invasive relatives (Gambusia sp.).Behavioral Ecology and Sociobiology 57:256-266.
- Relyea, R.A. 2001. Morphological and behavioral plasticity of larval anurans in response to different predators. Ecology 82: 523-540.
- Robinson, B.W. and D.S. Wilson 1998. Optimal foraging, specialization, and a solution to Liem's Paradox. American Naturalist 151: 223-235.

Salo, P., E. Korpimaki, P.B. Banks, M. Nordstrom and C.R. Dickman. 2007. Alien predators are more dangerous than native predators to prey populations. Proceedings of the Royal Society of London B 274:1237-1243.

- Scheiner, S.M. and M.R. Willig 2005. Developing unified theories in ecology as exemplified with diversity gradients. American Naturalist 166: 458-469.
- Schlaepfer, M. A., M. C. Runge, and P. W. Sherman. 2002. Ecological and evolutionary traps. Trends in Ecology & Evolution 17:474-480.
- Schmitz, O.J. 2004. Trophic cascades: the primacy of trait-mediated indirect interactions. Ecology Letters 7: 153-163.
- Schmitz, O.J., J.H. Grabowski, B.L. Peckarsky, E.L. Preisser, G.C. Trussell and J.R. Vonesh. 2008. From individuals to ecosystem function: Toward an integration of evolutionary and ecosystem ecology. Ecology 89: 2436-2445.
- Schoener, T.W. 1971. Theory of feeding strategies. Annual Review of Ecology and Systematics 2:369-404.
- Schoener, T. W. 1974. Resource partitioning in ecological communities. Science 185: 27-39.
- Schoener, T. W. 1989. The ecological niche. Pages 79-114 in J. M. Cherrett, editor. Ecological concepts: the contribution of ecology to an understanding of the natural world. Blackwell Scientific, Oxford, England.
- Sih, A. 1980. Optimal behavior: can foragers balance two conflicting demands? Science 210:1041-1042.
- Sih, A. 1987. Predators and prey lifestyles: an evolutionary and ecological overview. Pages 203-224 in W.C. Kerfoot and A. Sih, eds. Predation: direct and indirect impacts on aquatic communities. University Press of New England, Hanover, NH.
- Sih, A. and R.D. Moore. 1990. Interacting effects of predator and prey behavior in determining diets. Pages 771-796 in R.N. Hughes, ed. Behavioural mechanisms of food selection. NATO ASI Series. Springer-Verlag, Heidelberg, New York.
- Sih, A. 1992. Prey uncertainty and the balancing of antipredator and feeding needs. American Naturalist 139:1052-1069.
- Sih, A. 1998. Game theory and predator-prey response races. Pages 221-238 in L.A. Dugatkin and H.K. Reeve, eds. Game Theory and Animal Behavior. Oxford University Press, Oxford.

- Sih, A. and B. Christensen 2001. Optimal diet theory: when does it work, and when and why does it fail? Animal Behaviour 61:379-390.
- Sih, A., G. Englund and D. Wooster. 1998. Emergent impacts of multiple predators on prey. Trends in Ecology and Evolution 13:350-355.
- Stephens, D.W. 2002. Discrimination, discounting and impulsivity: a role for an informational constraint. Philosophical Transactions of the Royal Society 357:1527-1537.
- Stephens, D.W. and J.R. Krebs 1986. Foraging Theory. Princeton University Press, Princeton, NJ.
- Stephens, D.W. 2007. Models of information use. Pages 31-58 in D.W. Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Stephens, D.W., J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- R. W. Sterner and J. J. Elser. 2002. Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere, Princeton University Press, Princeton, NJ.
- Svanback, R. and D.I. Bolnick 2005. Intraspecific competition affects the strength of individual specialization: an optimal diet theory method. Evolutionary Ecology Research 7:993-1012.
- Svanback, R. and D.I. Bolnick. 2007. Intraspecific competition drives increased resource use diversity within a natural population. Proceedings of the Royal Society London B 274:839-844.
- Tilman, D. 1982. Resource Competition and Community Structure. Princeton University Press, Princeton, NJ
- Tollrian, R. and C.D. Harvell 1999. The Ecology and Evolution of Inducible Defenses. Princeton University Press, Princeton, NJ.
- Waite, T.A. and K.L. Field. 2007. Foraging with others: games social foragers play. Pages 331-362 in D.W. Stephens, J.S Brown and R.C. Ydenberg, eds. Foraging. Behavior and Ecology. University of Chicago Press, Chicago.
- Werner, E.E. 1977. Species packing and niche complementary in 3 sunfishes. American Naturalist 111: 553-578.
- Werner, E.E. and B.R. Anholt. 1993. Ecological consequences of the tradeoff between growth and mortality rates mediated by foraging activity. American Naturalist 142:242-272.
- Werner, E.E. and J.F. Gilliam 1984. The ontogenetic niche and species interactions in sizestructured populations. Annual Review of Ecology and Systematics 15:393-425.
- Werner, E.E., J.F. Gilliam, D.L Hall and G.G. Mittelbach. 1983. An experimental test of the effects of predation risk on habitat use in fish. Ecology 64:1540-1548.
- Werner, E.E. and D.J. Hall. 1974. Optimal foraging and size selection of prey by bluegill sunfish (*Lepomis macrochirus*). Ecology 5: 1042-1052.

Table 4.1. Summary of propositions of foraging theory.

Basic Propositions that foraging theory builds on

1. Foraging patterns maximize fitness or a correlate of fitness.

2. Foraging patterns depend on the range of options available to the forager and on how each available option affects fitness or a correlate of fitness.

3. Foraging behavior balances conflicting demands – tradeoffs are important in shaping foraging behavior.

Propositions about key factors influencing foraging behavior

4. Foraging behavior varies depending on the forager's other traits

5. Foraging behavior responds to environmental heterogeneity (including other species)

6. Foraging theory requires foragers to estimate parameters that influence the fitness associated with different foraging options.

Propositions that are major predictions of foraging theory

7: A forager should stay in a patch as long as its current reward rate (the 'marginal value' of the patch) is higher than the average reward rate for the rest of the habitat, and it should leave as soon as the current reward rate is no better than the expected rate elsewhere.

8: When high value prey are abundant, foragers should specialize on high value prey (i.e., reject low value prey), whereas when high value prey are scarce, foragers should have generalized diets (i.e., they should attack both low and high value prey when they are encountered).

9. With a sigmoid fitness function, foragers with low average feeding rates have little to lose, and should thus be risk prone while foragers with high average feeding rates have little to further gain, and should thus be risk averse.

10. Animals that have low assets (e.g., low energy reserves, or low reproductive value) should be bold (accept greater predation risk) to gain more energy, whereas animals with high assets should be cautious to protect their assets from catastrophic loss (e.g. death).

11. If foragers are free to move whenever they could do better elsewhere, then at the ideal (optimal) free distribution, all patches should yield equal reward rates.

12. Foraging behavior often balances the conflicting demands of gaining energy and avoiding predation risk.

Propositions that relate foraging theory community ecology

13: Foraging theory provides a mechanistic basis for understanding niches and patterns of competition that often underlie patterns of coexistence.

14: Predators often have strong non-consumptive effects on communities that are mediated by shifts in foraging behavior.



Figure 4.1. The predator-prey sequence beginning with an encounter, proceeding to consumption of prey. Predation rates depend on encounter rates and the sequence of probabilities shown in boxes. Predator and prey traits (including behaviors) that influence these rates and probabilities are shown above and below the boxes, respectively. Foraging theory focuses primarily on two predator behaviors – predator optimal patch use and optimal diets. The flowchart also shows some other predator and prey behaviors that potentially influence predation rates.



Figure 4.2. A graphical presentation of optimal patch use theory. Shown is the forager's net energy intake rate as a function of time spent in a patch. Due to prey depletion (or prey going into hiding), the forager's current net intake rate (the marginal value of the patch) should decline with time spent in a patch. The dashed horizontal line is the average net intake rate for the overall habitat. An optimal forager should stay in a current patch as long as its current net intake rate is higher than the average for the overall habitat. The optimal time to spend in a patch (T*) is longer in better patches.



Figure 4.3. An integrative overview of inter-relationships between foraging theory (which attempts to explain foraging behavior) and other major fields of ecology and evolution. Foraging behavior, shaped by natural selection, balances costs and benefits that depend on the forager's traits and state, as well as on social and species interactions and ecosystem properties. In turn, the resulting foraging behaviors are major components of resource/risk niches that underlie species interactions that explain many patterns in population, community and ecosystem ecology. All of this is influenced by environmental uncertainty.

Chapter 5: Ecological Niche Theory Jonathan M. Chase

Every species has a range of environmental parameters where it can have positive population growth; these parameters are said to be part of the species' niche (e.g., Hutchinson 1957). Further, we can distinguish among the fundamental niche, which consists of the range of environmental parameters where the species can maintain positive population in the absence of other species, and the realized niche, which consists of the range of environmental parameters where the species can maintain positive growth in the presence of other species (Hutchinson 1957). As such, niche theory pervades a majority of ecological investigation. Among other things, a species' niche is defined by evolutionary processes mold a species' traits as it interacts with its environment and other organisms, and influences the biogeographic range of species, the number of species that can coexist both locally and regionally, the strengths of interspecific interactions, the relative abundance of species (e.g., which are common and which are rare), and the role that species play in ecosystem functioning (Chase and Leibold 2003).

A species' ecological niche is roughly divided into two related components: (1) the suite of biotic and abiotic factors that influence a species ability to persist in a given locality, and (2) the impact that a species has on those factors (Leibold 1995, Chase and Leibold 2003). Importantly, I will explicitly incorporate the possible role of stochastic processes, both environmental (e.g., Chesson 2000) and demographic (e.g., Hubbell 2001) into the propositions of the niche theory described below. Such stochasticities are inherent to a more general niche theory, and have been investigated in concert with niche theory for some time (MacArthur and Wilson 1967, Chesson and Warner 1981, Strong et al. 1984, Tilman 2004). Thus, even though the neutral theory espoused by Hubbell (e.g., 2001) is not directly covered in this volume, my hope is to continue to move towards a more general synthesis of niche and neutral approaches as two sides of the same 'coin' (see also Gravel et al. 2006, Leibold and McPeek 2006, Adler et al. 2007).

AN ABBREVIATED HISTORY OF NICHE THEORY

Since its inception by Johnson (1908), through its more thorough development by Grinnell (e.g., 1917), Elton (1927), Gause (1934), among others (reviewed in Schoener 1989, 2008, Chase and

Leibold 2003), the niche concept has traversed a tumultuous road in the development of ecological enquiry. This is despite the niche being a consistently core ecological concept. Two main issues have plagued niche theory.

As a first plague, the niche concept has had multiple meanings in an ecological context, some of which are too vague or grandiose to be of use, and some of which are somewhat contradictory (see e.g., Schoener 1989, 2008, Chase and Leibold 2003). The duality of niche definitions has been often discussed (Schoener 1989, Leibold 1995); one view more focused on the 'place' that species live, including environmental and interspecific factors (e.g., Grinnell 1917, Hutchinson 1957), and the other focused on the 'role' of species in the community (e.g., Elton 1927). Added to this confusion is the fact that there are many related and overlapping terms in the literature, including habitat (Whittaker et al. 1973), which describes the environmental features in which species live (e.g., Whittaker et al. 1973, Southwood 1977), and guild, which describes a group of organisms with similar needs (Root 1967). Although these terms are often used, and can be useful, I avoid these terms and their associated concepts owing to their own inherent confusions and subjectivity.

The second plague is that niche theory has endured several rather contentious debates that interestingly seem to have arisen in ~20-30 year recurrences (see also Cooper 1993). In each case, the primary contention revolves around the predictability of communities and their matching to environmental conditions, ranging from the historical debate between the superorganismic versus individualistic organization of communities (e.g., Clements 1916 vs. Gleason 1926) to the contemporary debate between niche versus neutral theorists (e.g., Hubbell 2001 vs. Chase and Leibold 2003).

Although not formally designated in the context of niches, the debate between Clements (e.g., 1916) and Gleason (e.g., 1926) regarding the structure of communities revolved around how species in communities responded to environmental variation. Clements' view was more deterministic, suggesting that for any given set of environmental conditions, a specific community (the 'superorganism') would develop. Gleason's view, however, considered that species associations were less aligned, and that often community structure could be less predictable due to stochastic processes.

In 1957, the Cold Springs Harbor Symposium on Quantitative Biology convened a diverse group of demographers studying human and other animals, population and community

ecologists, and evolutionary biologists. Though the goals of this meeting were manifold, one recurrent theme was the debate about controls of populations, and whether they were more likely to be controlled by internal processes such as density dependence and species interactions or external processes such as weather (see also Andrewartha and Birch 1954 vs. Lack 1954). In his now classic 'Concluding Remarks' to that symposium, Hutchinson (1957) presented his quantitative definition of a species' niche as an 'n-dimensional hypervolume' of factors that influence the persistence of that species. He also differentiated among a species' 'fundamental niche', describing the factors where a species *could* possibly live in the absence of any other biotic or historical factors, and the 'realized niche', describing the factors that influence where a species actually does live in the presence of other interacting species (and also dispersal, history, and other constraints).

The third debate where the niche concept played a central role arose during the late 1970s and early 1980s. Following Hutchinson's (1957) definition of niche, his address regarding the limits to biodiversity (Hutchison 1959), and the works of several of his students and colleagues, most notably MacArthur (e.g., MacArthur 1958, 1960, 1964, 1970, MacArthur and Levins 1967), a renaissance of studies patterns of species coexistence and limits to niche similarity occurred (reviewed in Levin et al. 1972, MacArthur 1972, Vandermeer 1973, Schoener 1974). In a series of papers, Dan Simberloff and colleagues (e.g., Simberloff 1978, Connor and Simberloff 1979, Simberloff and Boecklen 1981, Simberloff 1983) challenged much of the empirical evidence for this theory on statistical grounds. Specifically, these authors developed a rigorous null model approach (reviewed in Gotelli and Graves 1996) to suggest that many of the observed patterns put forth to validate the niche theory based on competitive interactions and coexistence were not different from what would have been expected by random chance alone. Following a rather intense, and at times venomous, debate (see e.g., Strong et al. 1984), ecology became a more experimental and statistically robust science.

The ongoing debate regarding the importance of niche theory against its more stochastic alternatives intensified at the beginning of this decade with the publication of Hubbell's (2001) 'Unified neutral theory'. The neutral theory emerged from a conceptual realm related to the null model statistical approaches discussed in the preceding paragraph, along with a close allegiance with models of neutral processes in population genetic systems. Specifically, Hubbell's (2001) neutral model (see also MacArthur and Wilson 1967, Caswell 1976, Bell 2000) predicts patterns

of species diversity, composition, and relative species abundances based on stochastic processes alone, without the invocation of niche differences among species. In doing so, Hubbell's primary goal was to see how well such a 'neutral' theory can do in predicting patterns through space and time—and the answer was: often pretty well! The potential success of the neutral model begs the question of whether the premise of basing ecological models on niches is an appropriate starting place, or whether the neutral model is a better place to start (Volkov et al. 2003, 2007). As a result, there has been a rather contentious debate regarding the ability of the neutral model to predict patterns of community structure relative to more niche-based models (reviewed in Chave 2004, McGill 2006).

In each of the above cases, proponents of niche theory emphasize determinism and predictability, whereas the opponents suggested that other forces, primarily stochastic ones, often predominated. Obviously, then, a more complete niche theory needs to include both deterministic and stochastic elements. Hubbell's neutral theory was directly borrowed from concepts of stochastic dynamics in population genetics; ecological drift in communities is directly analogous to genetic drift in populations. Population geneticists recognize that natural selection and genetic drift act simultaneously, but differ in their relative importance as a result of several factors, including population size, dispersal, and the strengths of natural selection (Templeton 2006). Similarly, ecological drift and niche-selection act simultaneously in communities (Leibold and McPeek 2006, Adler et al. 2007), and the relative importance of the two processes should also vary depending on community size, dispersal, and the strengths of niche selection (Chase 2003, 2007). In the following sections, I will revisit niche theory from first principles, show how a variety of related theories collapse into a few general concepts, and theory.

DEFINING THE NICHE AND ITS PROPOSITIONS

Following Leibold (1995) and Chase and Leibold (2003), the niche of a species can be separated into two fundamental units; the requirement component and the impact component. Here, we define these components in the context of consumer-resource equations for simplicity of presentation and for logical progression (based on MacArthur 1972, Tilman 1982). However, this model presentation is not intended to be complete or universally applicable, and instead is

simply intended to serve as a backdrop for a more general formulation of the niche concept. In addition, it is important to note that this view of the niche is orthogonal to other more historical niche concepts which depicted a species' niche (and potential overlap with other species) as the number/types of resources on which they consumed (and often the degree to which they used them) (e.g., MacArthur and Levins 1967).

The *requirement* component of the niche denotes the minimum or maximum level of a particular factor where a species can persist in a given habitat. This can be depicted graphically by exploring how a species' birth and death rates vary with a factor that influences a species niche. For example, with increasing levels of a limiting resource, a species' birth rate should increase (Figure 5.1a); its death rate can be constant along this range, or could decrease if higher resources allows the organism to resist mortality from other sources (either way, the qualitative conclusions are the same). If the externally supplied availability of resources is such that births exceed deaths, a species can persist, whereas when the availability of resources are such that deaths exceed birth, a species does not persist.

Resources will be maintained at the equilibrium point when the consumer's birth and death rates are equal; this equilibrium is known as the R* of a consumer for a resource (Tilman 1982), and defines the lower boundary of the consumer's requirement niche for that resource. As a result, when the externally supplied level of resources are such that the birth rate is equal to or greater than the death rate, the species can persist. If, alternatively, the limiting factor of interest is a predator, a species can persist where the level of predation is such that births equals deaths, and is defined as a species' P* (Holt et al. 1994). In fact, any such factor that influences births and/or deaths, including mutualists, environmental stressors, and/or spatio-temporal heterogeneities, can be determined by calculating the equilibrium where birth rates equal death rates.

The boundaries of the requirement component of the niche can be extended to any number of niche factors, though generally, only two dimensions are graphically depicted. For example, the requirement component can be extended to two resources, and graphically depicted by plotting the relationship through the two-dimensional state space—the combination of factors where the species birth rate is equal to the death rate (equilibrium) is termed its' zero net growth isocline (ZNGI) (Figure 5.1b). ZNGIs such as that depicted in Figure 5.1b can in actuality be derived for any combination of factors that influence the population growth (births and deaths) of

an organism, either positively or negatively. The well-studied factors include predators, stressors, and spatio-temporal heterogeneities (Holt et al. 1994, Leibold 1996, 1998, Wootton 1998, Chase and Leibold 2003), but other factors, including those that can have positive effects, such as ecosystem engineers and mutualists can also be relatively easily incorporated into this framework (Chase and Leibold 2003).

The *impact* component of a species' niche denotes the influence the species has on the niche factor of interest. For example, if the factor is a resource, the consumer will act to drive the level of that resource down towards the R*, as long as the external supply of that resource exceeds the R* (Figure 5.1a). Alternatively, if the factor of interest is a predator, the species will act to drive the density of the predator higher, and thus the impact vector will point upwards along the predator axis. Finally, if the factor of interest alters a species' birth and death rates, but the species has no influence on that factor, such as externally imposed stressors (e.g., disturbance), the species will have no impact on that factor.

As with requirements, the impacts of a species can be depicted in two (or more) dimensions for any two limiting factors by taking the vector sum of the impacts on each factor. For example, if there is a consumer of two resources, the consuming species has a negative impact on both resources, and the sum of those two (horizontal and vertical) vectors represents the impact of the species on the two resource system (Figure 5.1b). The slope of the vector indicates the relative magnitude of the impact of the species on each resource; if the species has equal effect on both resources, the slope will be one, whereas it will be steep or shallow, if the consumer of interest has a greater impact on one resource relative to the other. As above, impact vectors can be drawn for any two-dimensional system, including combinations of resources (negative effects), predators and mutualists (positive effects), and environmental stressors (no effects).

Finally, in order to determine whether a species can actually persist in a given locality, we need to consider the external *supply* of the niche factor of interest. Depending on the nature of the niche factor, the supply can be a rate denoted by a vector, or as a single point in the state space. If the factor is a resource with an inherent growth rate (e.g., logistic growth towards a carrying capacity), the supply vector will point towards the carrying capacity of that resource in the absence of its consumer. Likewise, the supply vector will point towards an externally supplied rate of input in the case of abiotic resources such as light or water. If, alternatively, the

resource is in mass balance (sensu Holt et al. 1994), as is the case with many limiting nutrients, the supply can be depicted as a point in the state space where the resource would be in the absence of any consumers. External 'supplies' of niche factors other than resources can also be depicted in a similar manner. For example, if the niche factor is some stress like a disturbance, this can be depicted as a point in the state space showing the rate or intensity of that disturbance. Finally, if the niche factor of interest is dependent on the species being modeled, for example, if the niche factor is a predator that can not persist without the prey being modeled, its external supply will be zero.

Just as in standard predator-prey theory (see Holt Chapter 7), in the simplest scenario, a species will maintain its resource at its R* regardless of the external supply of that factor. Thus, variation in the standing levels of resources does not necessarily imply variation in the supply rates of resources, but rather, can indicate differences in the R* of the consumers in a system. Adding a factor that increases the death rate of a species (or decreases its birth rate), such as a predator, will increase the equilibrium abundance of the resource left behind in the system (e.g., a trophic cascade) (Holt et al. 1994, Leibold 1996). Niche factors on which a species does not have an impact (e.g., a stressor such as disturbance) are not influenced by the presence of that species (Wootton 1998, Chase and Leibold 2003).

PROPOSITIONS OF THE NICHE THEORY

Armed with these conceptual niche components, we can establish a series of propositions (*sensu* Scheiner and Willig 2008, Chapter 1) that a synthetic niche theory can provide (Table 5.1). Importantly, this synthetic niche theory can directly incorporate deviations from the simplest predictions described above. It can do this by including complexities in species physiological, life-historical and behavioral attributes (Tilman 1988, Vincent et al. 1996, Chase 1999, Chase and Leibold 2003), species interactions in food webs (Holt et al. 1994, Leibold 1996, 1998, Chase and Leibold 2003), a variety of spatial and temporal environmental heterogeneities (Tilman and Pacala 1993, Chesson 2000, Chase and Leibold 2003), as well as the stochasticity inherent to the neutral theory (Chesson 2000, Adler et al. 2007). Many of the details of this synthetic niche theory and the propositions discussed below are only superficially covered here, and for more depth, see Chase and Leibold (2003) (see also MacArthur 1972, Tilman 1982,

1988, Holt et al. 1994, Leibold 1995, 1996, Grover 1997, Chesson 2000, Adler et al. 2007, among others).

Proposition 1

When there is a single limiting factor, the species that can maintain that factor at the lowest (or highest) level will outcompete all others. When there is only a single limiting factor (and no spatio-temporal heterogeneities in that factor), only one species will exist at equilibrium. If the factor is a resource, the species that exists alone will be the one that can persist on the lowest level of that resource (lowest R*). This basic proposition is often know as Tilman's R* rule (Tilman 1976, 1982), although its roots can be seen earlier (e.g., MacArthur 1972). The derivation of this principle is relatively straightforward. Assuming one species maintains the resource at its R* owing to the consumer-resource dynamic, the level of that resource will be too low for any other species to invade. As alluded to above, the same principle can be derived from the standard Lotka-Volterra predator-prey model (see Holt Chapter 7), where the vertical isocline of a predator that is closest the abscissa will outcompete any predator with an isocline further from the abscissa.

This basic principle can be applied to any type of limiting factor that influences a species' birth and/or death rates. For example, if the factor of interest is a predator instead of a resource, the P* rule can be obtained (e.g., Holt et al. 1994). In this case, through apparent competition for a shared predator, the species that is able to maintain positive population growth rate on the highest level of predators (highest P*) will outcompete any others, and exist alone. Alternatively, if the factor is a stressor that influences birth and/or death rates, the species that will persist alone will depend on the level of the stress, since species are assumed to have no impact on stress; the more stress tolerant species will persist alone at high stress, and the stress intolerant species will persist alone at lower stress (Wootton 1998, Chase and Leibold 2003)

If there are multiple limiting factors, but only one species is superior on all of those factors such that it is a 'Hutchinsonian Demon' (*sensu* Kneitel and Chase 2004), it will outcompete all others and exist alone. Instead, in order for more than one species to be able to coexist locally, there must be more than one limiting resource, and those species must trade-off in their ability to utilize those resources, which leads us to the next proposition

Proposition 2

For more than one species to coexist locally, they must trade-off in their relative abilities to compete for different limiting factors, as well as in their relative impacts on those limiting factors. For two species to coexist locally in a spatio-temporally homogeneous environment when there are two limiting resources, they must trade-off in their ability to utilize those resources. That is, for potential coexistence, the ZNGIs of the two species must intersect, where one species has a lower R* for one resource, and the other a lower R* for the other resource. The same arguments can be extended to multiple resources, so long as species show perfect trade-offs for those resources (e.g., each species has its lowest R* on a different resource) (Levin 1970, Kneitel and Chase 2004). While such trade-offs are necessary for local coexistence, this trade-off alone does not guarantee local coexistence. First, there must be a second trade-off in the impacts of those species on the resources in order for the equilibrium to be locally stable. Specifically, the impacts of the two species need to be such that each species has a greater relative impact on the resource that it finds most limiting (has a higher R*). That is, if a species has traits that allow it to persist at relatively low levels of a particular resource (low R^*), those traits need to trade-off with other traits that allow it to impact (e.g., consume) the resource that it finds more limiting (higher R*). This trade-off allows intraspecific effects of each species on itself (through their impacts on the resource) to be greater than the interspecific effects of each species on the other species, which is the standard criteria necessary for coexistence in any multispecies community.

In addition to these coupled trade-offs in requirements and impacts among the competing species local coexistence also requires particular conditions available in the environment. First, the supply rate of the resources must exceed that of the R*s for the species. Second, the ratios of the supplies of those resources needs to be intermediate, and not skewed strongly towards any one resource (in which case, the species will the lowest R* for that resource can outcompete the others through quorum effects).

This emphasizes that a fundamental feature of niche differences and coexistence is not just differences in the traits of species, but in the characteristics of the environment in which the organisms live. Thus, environmental context is an essential component of understanding niche theory and coexistence among species, although this has not always been explicitly recognized.

These general principles of local coexistence are true for any species that interacts with two or more factors which influence their requirement niche (e.g., birth and death rates) and, and on which they have impacts. For example, when two prey species share two predators, they can potentially coexist if: (1) they trade-off in their P*s, such that each prey species is least influenced by, and thus a better apparent competitor (higher P*) for, one of the predator species; and (2) each species has a bigger positive impact (e.g., is a better food source) on the predator for which it is a worse apparent competitor (lower P*) (Leibold 1998). Likewise, when two prey species share a common resource and a common predator, they will coexist when the species that is a superior resource competitor is simultaneously more susceptible to predation than the weaker resource competitor (i.e., a keystone predator), and the better resource competitor (lower R*) is also better food (higher impact) for the predator (Holt et al. 1994, Leibold 1996, Chase et al. 2000). Importantly, if the species do not impact one or more of the two limiting resources, local coexistence is not possible, but regional coexistence is (see below).

Proposition 3

Temporal variability can allow species to coexist that otherwise could not. An extremely important, but sometimes overlooked mechanism by which species can coexist locally is temporal variability in the availability of resources through time. If the supply rates of resources are variable, more than one species can coexist, even on a single resource, if they trade off in their abilities to utilize the resource when it fluctuations at different availabilities. For example, rainfall and temperatures vary seasonally, and species can coexist locally by partitioning their competitive abilities across seasons (e.g., Hutchinson 1961, Grover 1997). Additionally, variability can emerge among years, for example, when precipitation varies inter-annually in arid systems. In this case, if some species are favored under wetter years, and others under drier years, they can coexist locally so long as the time frame of the variability in rainfall is not longer than the time it takes for the species to go locally extinct (i.e., the storage effect; Chesson and Warner 1981, Chesson 2000). Similarly, species can coexist if they trade-off in their relative abilities to compete for resources and in their ability to colonize habitats following temporal disturbances; the colonization-competition trade-off (Hastings 1980, Tilman 1994). Finally, temporal variability can be internally driven by the dynamics of the consumers themselves (e.g., when responses are non-linear), and does not have to be externally driven. If consumers create a

non-equilibrial (cycling) dynamic in their resource base, they can coexist by specializing on different densities of their prey as it cycles (Armstrong and McGhee 1976, Huisman and Weissing 1999, Abrams and Holt 2002).

Proposition 4

For two or more species to coexist regionally, they must trade-off in their relative abilities to compete for different limiting factors, and those resources must be variably abundant in space. Even when species show the appropriate trade-offs for local coexistence, the supplies of resources are often skewed such that one species dominates in any given locality (i.e., no local coexistence). However, if the supply rates of those resources vary spatially (e.g., along a nutrient gradient), the species can coexist regionally if the availability of those resources varies spatially, so long as the species show perfect trade-offs in their usage of those two resources (Tilman and Pacala 1993, Chase and Leibold 2003). As with temporal variability, spatial variability in resources can also be generated by the organisms themselves, allowing species to coexist (Wilson et al. 1999, Wilson and Abrams 2005). Finally, spatial heterogeneity combined with dispersal can allow species to coexist both locally and regionally through source-sink processes, whereas higher dispersal rates can reduce local and regional diversity by allowing the better colonizers to 'swamp out' the better competitors (Amarasekare and Nisbet 2001, Mouquet and Loreau 2003).

Importantly, the temporal and spatial heterogeneities discussed in this and the above proposition exposes a critical misrepresentation of niche theory by proponents of a purely neutral theory (Hubbell 2001, Bell 2003). Specifically, neutral theory proponents have suggested that niche theory is predicated on the fact that the numbers of resources determine the number of species that can coexist (N resources leads to N species; Levin 1970), and that this can not account for the very high levels of diversity often observed, particularly in tropical areas. However, simple temporal and spatial heterogeneities, even when there are only a few limiting resources, can allow very large numbers of species to coexist (Tilman 1994, 2004).

Proposition 5

Stochastic ecological drift will counter-balance deterministic, niche-based, processes when community size is small. Though niche and neutral theories have been treated as dichotomous

alternatives (e.g., McGill et al. 2006), the primary components of each theory, determinism and stochasticity, occur simultaneously (Gravel et al. 2006, Leibold and McPeek 2006, Adler et al. 2007, Chase 2007). Just as the stochasticity associated with genetic drift can interact with the determinism associated with natural selection in population genetics, stochastic ecological drift interacts with niche-selection (e.g., the environmental filtering of the species pool) in communities (Adler et al. 2007, Chase 2007).

The relative importance of stochasticity in community assembly will be highly contingent on both local and regional factors. In direct analogy to small population sizes increasing the relative importance of genetic drift in population genetics, small community sizes (individuals and species) increase the relative importance of stochastic ecological processes (Drake 1991, Chase 2003, Fukami 2004, Orrock and Fletcher 2005). This results from simple probability theory, where the possibility that stochastic processes can lead to large influences on overall community structure increases with smaller communities (and thus fewer individuals/species). Likewise, for analogous probabilistic reasons, the relative importance of stochasticity will be reduced among communities that are more strongly connected by dispersal than in communities that are isolated (Forbes and Chase 2002, Chase and Ryberg 2004, Shurin et al. 2004).

Proposition 6

Stochastic ecological drift will counter-balance deterministic, niche-based, processes when niche-selective forces are weak. The relative strength of niche-selection on the pool of species that can colonize a given locality will also importantly influence the relative role of stochasticity in the assembly of communities. Stochastic processes will be most important when environmental filters (niche selection) on the species pool are weak, such that the 'realized' pool of species that can colonize a given locality is larger. When species pools are larger, stochastic processes will play a more important role in the assembly of local communities by definition (Chase 2003, 2007). For example, stochasticity in the assembly of communities, and thus the variation in community membership, appears to be more important in higher productivity than in lower productivity systems, and this may result because a larger number of species in the species pool can persist in higher productivity (Inouye and Tilman 1995, Chase and Leibold 2002, Chase 2003, Chalcraft et al. 2008). Likewise, communities in more disturbed or stressful environments appear to be less variable in community composition than less disturbed/stressful environments,

also likely as a result of disturbance/stress filtering out certain members of the broader species pool (Booth and Larson 1999, Chase 2003, 2007, Trexler et al. 2005, Jiang and Patel 2008, Lepori and Malmqvist 2008).

RELATIONSHIP TO OTHER THEORIES IN ECOLOGY

Because the niche generally describes the distribution and abundance of a given organism, as well as the factors that influence coexistence (diversity) among organisms, the domain of niche theory is rather inclusive of much of the domain of ecology as a whole. As a result of this broad association, it also has close ties with several of the other constituent theories in ecology.

- Foraging theory (Sih Chapter 4) defines the ways in which organisms acquire the resources they need, the evolutionary causes of such acquisition, and the behavioral mechanisms by which they do so. As resource acquisition is fundamental for an organism to persist within its defined niche, foraging theory represents a primary mechanism by which organisms are able to achieve their requirements and have their impacts on their resources. For example, an important tool in foraging ecology is measuring the density of resources in a patch on which a forager 'gives up', and moves to another foraging patch (e.g., Brown 1988). This same tool can be used to understand the relative position of a species' R* in an environment, as well as how various factors (e.g., predation risk) influence that R* (Chase and Leibold 2003).
- Population dynamics theory (Hastings Chapter 6) defines the ways in which an organism's population grow in the context of their own birth and death rates, and thus has obvious connections to the births and death rates that go into defining the parameters of the niche. Further, interspecific interactions, including predator-prey interactions (Holt Chapter 7), alter these birth and death rates, thus directly alter the structure of the organisms' niche.
- Patterns of community structure through time, as predicted by succession theory (Pickett et al. Chapter 9), and in space, as predicted by both metacommunity theory (Leibold Chapter 8) and environmental gradient theory (Fox et al. Chapter 13), depend on the interactions among species and their spatio-environmental context. As such, species' niches fundamentally underlie the mechanisms leading to these patterns.

- Ecosystems theory (Burke and Lauenroth Chapter 11) examines the fluxes of energy and nutrients through organisms and their environment. In many circumstances, the traits of species, as defined by their niches, as well as the diversity and composition of species in a community, influence the importance of their roles in ecosystems.
- A recent application of niche theory is 'species distribution modeling' (often called 'niche modeling'), which correlates spatio-temporal availability of broad-scale environmental features with species occurrences (e.g., Stockwell 2006). While species distribution modeling, by being correlational between distribution and environment, is rather simplistic relative to the aspects of niche theory discussed here, it has potential to link small- and large-scale processes. Further, such distribution modeling is often used to try to understand how species distributions might be altered with global change (e.g., Peterson et al. 2002), and thus provides important linkages with global change theory (Peters et al. Chapter 12).

CONCLUSIONS

Despite repeated attempts over many decades (Grinnell 1917, Elton 1927, Hutchinson 1957, MacArthur 1972, Tilman 1982, Chase and Leibold 2003), niche theory has still not fully reached its potential as a central organizing theory of ecological enquiry. This is in part due to the fact that the niche is such a hard to define concept. Indeed, any definition, verbal or mathematical, of the niche can only represent a caricature of the complexity of any species' multivariate interactions with its environment and other species. Additionally, niche theory has not traditionally incorporated a variety of complexities, most notably spatio-temporal heterogeneities and stochasticity, which can often play an overriding role in determining the structure of natural communities. An emerging more general niche theory explicitly incorporates spatio-temporal heterogeneities (Tilman 1994, 2004, Chesson 2000) as well as the stochastic processes inherent to neutral theory (e.g., Leibold and McPeek 2006, Adler et al. 2007, Chase 2007). Such a synthetic theory will provide the groundwork for a more general theory of ecology as a whole.

Literature Cited

- Abrams, P. A. and R. D. Holt. 2002. The impact of consumer-resource cycles on the coexistence of competing consumers. Theoretical Population Biology 62: 281-295.
- Adler, P. B. and J. HilleRisLambers, and J. M. Levine. 2007. A niche for neutrality. Ecology Letters 10:95–104
- Amarasekare, P. and R. M. Nisbet. 2001. Spatial heterogeneity, source–sink dynamics, and the local coexistence of competing species. American Naturalist 158: 572–584.
- Andrewartha, H. G., and L. C. Birch. 1954. The Distribution and Abundance of Animals. University of Chicago Press, Chicago, Illinois.
- Armstrong, R. and R. McGehee. 1976. Coexistence of two competitors on one resource. Journal of Theoretical Biology 56: 449-502.
- Bell, G. 2000. The distribution of abundance in neutral communities. American Naturalist 155: 606–617.
- Bell, G. 2003. The interpretation of biological surveys. Proceedings of the Royal Society of London (B). 270: 2531-2542.
- Booth, B. D. and D. W. Larson. 1999. Impact of language, history, and choice of system on the study of assembly rules. In: Ecological Assembly Rules. E. Weiher and P. Keddy (eds). Cambridge University Press, pp. 206-229.
- Brown, J. S. 1988. Patch use as an indicator of habitat preference, predation risk, and competition. Behavioral Ecology and Sociobiology 22: 37-47.
- Caswell, H. 1976. Community structure neutral model analysis. Ecological Monographs 46: 327–354.
- Chalcraft, D. R., S. B. Cox, C. Clark, E. E. Cleland, K. N. Suding, Weiher E, and D. Pennington. 2008. Scale-dependent responses of plant biodiversity to nitrogen enrichment. Ecology 89: 2165-2171.
- Chase, J. M. 1999. Food web effects of prey size-refugia: variable interactions and alternative stable equilibria. American Naturalist 154:559-570.
- Chase, J. M. 2003. Community assembly: when does history matter? Oecologia 136: 489-498.
- Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. Proceedings of the National Academy of Sciences 104: 17430-17434.

- Chase, J. M. and M. A. Leibold. 2002. Spatial scale dictates the productivity-diversity relationship. Nature 415: 427-430.
- Chase, J. M and M. A. Leibold. 2003. Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press
- Chase, J. M. and W. A. Ryberg. 2004. Connectivity, scale dependence, and the productivitydiversity relationship. Ecology Letters 7: 676-683.
- Chave, J. 2004. Neutral theory and community ecology. Ecology Letters 7:241–253.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual Reviews of Ecology and Systematics 31:343–366.
- Chesson, P. L. and R. R. Warner. 1981. Environmental variability promotes coexistence in lottery competitive systems. American Naturalist 117: 923-942.
- Clements, F. E. 1916. Plant succession: An analysis of the development of vegetation. Carnegie Institute of Washington.
- Cold Spring Harbor Symposium on Quantitative Biology 1957. Vol. XXII
- Connor, E. F. and D. Simberloff. 1979. The assembly of species communities: chance or competition? Ecology 60: 1132-1140.
- Cooper, G. 1993. The competition controversy in community ecology. Biology & Philosophy 8: 359–384.
- Drake, J. A. 1991. Community-assembly mechanics and the structure of an experimental species ensemble. American Naturalist 137: 1–26.
- Elton. C. S. 1927. Animal Ecology. Sidgwick and Jackson, London
- Forbes, A. E. and J. M. Chase. 2002. The role of habitat connectivity and landscape geometry in experimental zooplankton metacommunities. Oikos 96: 433-440.
- Fukami, T. 2004: Assembly history interacts with ecosystem size to influence species diversity. Ecology 85: 3234-3242.
- Gause, G. F. 1934. The struggle for existence. Williams and Wilkins, London, England.
- Gleason, H. A. 1926. The individualistic concept of the plant association. Bulletin of the Torrey Botanical Club 53: 7-26.
- Gotelli, N. J. and G. R. Graves. 1996. Null Models in Ecology. Smithsonian Institution Press, Washington, DC.

- Gravel, D., C. D. Canham, M. Beaudet, C. Messier. 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecology Letters 9: 399-409
- Grinnell, J. 1917. The niche-relations of the California Thrasher. Auk 34:427-433.
- Grover J. P. 1997. Resource Competition. Chapman & Hall, New York. Hastings 1980
- Hastings, A. 1980. Disturbance, coexistence, history, and competition for space. Theoretical Population Biology 18: 363–373.
- Holt, R. D., J. Grover and D. Tilman 1994. Simple rules for interspecific dominance in systems with exploitative and apparent competition. American Naturalist 144: 741–771.
- Hubbell, S. P. 2001. The unified neutral theory of species abundance and diversity. Princeton University Press, Princeton.
- Huisman, J. and F. J. Weissing. 1999. Biodiversity of plankton by species oscillations and chaos. Nature 402:407-410,
- Hutchinson, G. E. 1957. Concluding remarks- Cold Spring Harbor Symposia on Quantitative Biology. 22:415-427.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia or Why are there so many kinds of animals? American Naturalist 93:145-159.
- Hutchinson, G. E. 1961. The paradox of the plankton. American Naturalist 95:137-140.
- Jiang, L. and S. N. Patel. 2008. Community assembly in the presence of disturbance: a microcosm experiment. Ecology 89: 1931-1940.
- Kneitel, J. M. and J. M. Chase. 2004. Trade-offs and community ecology: linking spatial scales and species coexistence. Ecology Letters 7: 69-80
- Lack, D. 1954. The natural regulation of animal numbers. Oxford University Press, Oxford.
- Leibold, M. A., 1995. The niche concept revisited: mechanistic models and community context. Ecology 76:1371-1382
- Leibold, M. A. 1996. A graphical model of keystone predation: effects of productivity on abundance, incidence and ecological diversity in communities. American Naturalist 147:784-812.
- Leibold, M. A. 1998. Similarity and local co-existence of species in regional biotas. Evolutionary Ecology 12: 95–110.
- Leibold, M. A. and M. A. McPeek. 2006. Coexistence of the niche and neutral perspectives in community ecology. Ecology 87: 1399–1410.

- Lepori, F. and B. Malmqvist. 2008. Deterministic control on community assembly peaks at intermediate levels of disturbance. Oikos 118: 471-479.
- Levin, S. A. 1970. Community equilibria and stability, and an extension of the competitive exclusion principle. American Naturalist 104: 413-423.
- MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. Ecology 39: 599-619
- MacArthur, R. H. 1960. On the relative abundance of species. American Naturalist 94: 25-36.
- MacArthur, R. H. 1964. Environmental factors affecting bird species diversity. American Naturalist 98: 387-397.
- MacArthur, R. H. 1970. Species packing and competitive equilibrium for many species. Theoretical Population Biology 1: 1-11.
- MacArthur, R. H. 1972. Geographical Ecology. 269pp. Harper and Row.
- MacArthur, R. H. and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, New Jersey.
- MacArthur, R. H. and R. Levins. 1967. The limiting similarity, convergence and divergence of coexisting species. American Naturalist 101: 377-385,
- McGill, B. J., B. A. Maurer, M. D. Weiser. 2006. Empirical evaluation of neutral theory. Ecology 87: 1411–1423
- Mouquet, N. and M. Loreau 2003. Community patterns in source–sink metacommunities. American Naturalist 162: 544–557.
- Orrock, J. L. and R. J. Fletcher, Jr. 2005. Changes in community size affect the outcome of competition. American Naturalist 166: 107-111.
- Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sánchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B.. 2002. Future projections for Mexican faunas under global climatic change scenarios. Nature 416: 626–629.
- Scheiner, S. M. and M. R. Willig 2008. A general theory of ecology. 1: 21-28.
- Schoener, T. W. 1974. Resource partitioning in ecological communities. Science 185: 27-39.
- Schoener, T. W. 1989. The ecological niche. Pages 79-114 in J. M. Cherrett, editor. Ecological concepts: the contribution of ecology to an understanding of the natural world. Blackwell Scientific, Oxford, England.
- Schoener, T. W. 2008. Ecological niche. Encyclopedia of Ecology. In press.

- Shurin, J. B., P. Amarasekare, J. M. Chase, R. D. Holt, M. F. Hoopes, and M. A. Leibold. 2004. Alternative stable states and regional community structure. Journal of Theoretical Biology 227: 359-368
- Simberloff, D. 1978. Using island biogeographic distributions to determine if colonization is stochastic. American Naturalist 112: 713-726.
- Simberloff, D. 1983. Competition theory, hypothesis testing, and other community ecological buzzwords. American Naturalists 122: 626-635.
- Simberloff, D. and W. Boecklen. 1981. Santa Rosalia reconsidered: Size ratios and competition. Evolution 35: 1206-1228.
- Stockwell, D. 2006. Niche modeling: Predictions from statistical distributions. Chapman & Hall/CRC.
- Strong, D.R. Jr, D. Simberloff, L. G. Abele, A. B. Thistle. 1984. Ecological Communities: Conceptual Issues and the Evidence. Princeton Univ. Press, Princeton, NJ.
- Templeton, A. R. 2006. Population genetics and microevolutionary theory. Wiley.
- Tilman, D. 1982. Resource competition and community structure. Princeton University Press, Princeton, New Jersey, USA.
- Tilman, D. 1988. Plant strategies and the dynamics and structure of plant communities. Princeton University Press, Princeton, New Jersey, USA.
- Tilman, D. 1994. Competition and biodiversity in spatially structured habitats. Ecology 75:2-16.
- Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences 101:10854-10861.
- Tilman, D. and S. Pacala. 1993. The maintenance of species richness in plant communities. Pp. 13-25 in Ricklefs, R. and D. Schluter, Eds., Species Diversity in Ecological Communities, University of Chicago Press, Chicago.
- Trexler, J. C., W. F. Loftus, and S. Perry. 2005. Disturbance frequency and community structure in a twenty-five year intervention study. Oecologia 145:140-152
- Vandermeer, J.H. 1972. Niche theory. Annual Review of Ecology and Systematics 3: 107-132.
- Vincent, T. L. S., D. Scheel, J. S. Brown, and T. L. Vincent. 1996. Trade-offs and coexistence in consumer-resource models: It all depends on what and where you eat. American Naturalist 148: 1038-1058.

- Volkov, I., J. R. Banavar, S. P. Hubbell, and Maritan. 2003. Neutral theory and relative species abundance in ecology. Nature 424: 1035–1037.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and Maritan. 2007. Patterns of relative species abundance in rainforests and coral reefs. Nature 450: 45-49.
- Wilson, W.G., C. W. Osenberg, R. J. Schmitt, and R. Nisbet. 1999. Complementary foraging behavior allows coexistence of two grazers. Ecology 80: 2358-2372.
- Wilson, W. G. and P. A. Abrams. 2005. Coexistence of cycling consumer species having localized interactions; Armstrong and McGehee in space. American Naturalist. 193-205.
- Wootton, J. T., 1998. Effects of disturbance on species diversity: a multi- trophic perspective. American Naturalist 152:803-825.

Table 5.1. The background and propositions that constitute the theory of niches

- 1. When there is a single limiting factor, the species that can maintain that factor at the lowest (or highest) level will outcompete all others.
- 2. For more than one species to coexist locally, they must trade-off in their relative abilities to compete for different limiting factors, as well as in their relative impacts on those limiting factors.
- 3. Temporal variability can allow species to coexist that otherwise could not
- 4. For two or more species to coexist regionally, they must trade-off in their relative abilities to compete for different limiting factors, and those resources must be variably abundant in space.
- 5. Stochastic ecological drift will counter-balance deterministic, niche-based, processes when community size is small.
- 6. Stochastic ecological drift will counter-balance deterministic, niche-based, processes when niche-selective forces are weak.



Figure 5.1. *A*. An example of how examining the responses of a species' birth and death rates to a factor of interest—resources in this case—can determine the requirement component of the niche. The intersection of birth and death rates indicates the level of resources necessary in order for the species to be at equilibrium (neither growing nor declining), and is denoted as a species' R*. The impact of a species on this resource is denoted by the arrow, which indicates the strength of the consumption effect of the species on the resource. *B*. The requirement component of a species for two resources. Note, here, the Zero Net Growth Isocline (ZNGI) denotes the combination of resources where birth rates equal death rates (equilibrium). Here, the impact vector is the sum of the species impacts on the two resources. The R*, ZNGI, and impact vector have underlying mathematical definitions (see Chase and Leibold 2003), but have been popularized because of their graphical representation. In addition, similar graphical (and analytical) forms can be derived for any niche factor that influences a species' birth and death rates.

Chapter 6: Single Species Population Dynamics and its Theoretical Underpinnings Alan Hastings

A theory of single species population dynamics is one of the most basic and oldest parts of ecology (Kingsland 1995) and much of applied ecology is based on it. Management of fisheries (Clark 1990) has often been based on single species approaches. Approaches to conservation, even those embodied in the endangered species act, are very much focused on single species (Lande et al. 2003). Developing appropriate approaches for these and other management questions certainly depends on the development of a theory of population dynamics. And, a deep understanding of the assumptions underlying this theory is certainly important in judging how and when this theory should be applied.

Not surprisingly, single species population theory has a particularly long history (Kingsland 1995), which some have argued goes back at least as far as the work of Fibonnacci (who lived from roughly 1170-1250), who developed his description of population dynamics in the form of the Fibonnacci series. Jumping ahead, the work of Malthus (1798) over 200 years ago can be thought of as a direct precursor to some of the work that will be presented here.

Malthus presented his work verbally, but this still contained much of what was later expressed mathematically. His suppositions that populations grew geometrically while food supplies grew arithmetically is particularly relevant to the material below. This approach is one that essentially starts from the simplest description of the dynamics of a population and then proceeds to ask why this simple description does not hold.

On the other hand, there has been substantial recognition of the importance of space and heterogeneity (Levin 1992) and stochasticity (Lande et al. 2003) and other factors for understanding the dynamics of populations. This begins to raise the central question of the level of appropriate detail to use in a description of the dynamics of a population. This is a difficult question, and obviously depends on the purpose for studying the dynamics of the population. One could start from the simple approach essentially contained in the work of Malthus and build up to this more complete description. I will present this approach below. Alternatively, one could start from as complete a description of single species population dynamics as possible and consider simplifications that would potentially lead to more tractable models and descriptions. This approach, too, has its merit, and I will also present this alternate approach.

Thus, the remainder of this chapter is organized as follows. I will first present the development which starts from a relatively complete microscopic description, and proceeds with looking at appropriate averaging to yield more concise and useful models. Essentially this approach depends on two parts. The first is clearly in developing as complete a description as possible. The second is in the choice of which aspects will be averaged over. Various implications and consequences of this approach will be considered. I will refer to this as the *microscopic approach*.

Then, the alternate approach which starts from the simplest possible descriptions of population growth will be considered. This development starting with an idealized description will be called the *macroscopic approach*. Here, rather than the next steps after presenting the first model being removing parts of the description, the next steps will be to include more and more biological detail. In this approach, the biological questions and issues will be clear, but as the discussion below will indicate, it is not always clear how the models and descriptions are developed and related to underlying mechanisms.

In some senses, there are analogues to these two approaches in the context of physical theories. One could describe the behavior of a gas by starting with idealized gas laws (expressed only in terms of bulk measures like pressure, temperature or volume) and then subsequently add complications that recognize specific properties of either the gas or situation being examined. Alternatively, one could start with a description of the motion of each individual gas molecule based on mechanics and average appropriately to determine behavior at a larger scale. Because even single species population dynamics involves much more heterogeneity and complexity than the dynamics of a gas, the task for the population biologist will be much more difficult.

Although implications and lessons from each development will be developed in the discussion of each development, the metapopulation approach (Hanski 1999) has been so important that it will be considered separately. The discussion of metapopulations will be brief and the main goal will be to indicate how ideas from the approaches here might inform the use and development of metapopulation theory. Connections to other levels, such as behavioral ecology or community ecology will also be briefly considered. The subject of single species population dynamics is extremely large, so all of the discussion here will be limited. Notably, little attention will be given to specific models for populations, as the arguments developed here will not depend on the details of particular models.

DOMAIN OF THE THEORIES

The domain of the theory of single species population dynamics is the understanding and prediction of the numbers of individuals in a single species. At this level, the domain is the same for either approach. Within this context, the focus could be on total numbers of individuals, but could also be on a description of a population that is structured.

Implicit in this choice of domain is the underlying assumption that the dynamics of a single species can be understood without consideration of other species. Clearly, the extent to which this is useful depends strongly on the question asked, the purpose for the model. Single species models are likely to be relatively poor tools for understanding biodiversity. The extent to which these models are useful for managing species, either for exploitation (fisheries), or to control undesirable species, will depend very much on the specific system.

MICROSCOPIC APPROACH

I discuss the theory of single species population dynamics using two different, but complementary approaches. The first approach is the one that is based on simplifying a description that is more or less complete, at least at the single species level. The idea here is that this top down approach has two advantages. First, a very complete description of the domain of single species ecology will make it very clear what is being left out, and at least suggest how the omitted factors might be important. Second, the specification of a relatively complete model makes clear what is being left out in a particular instance.

I begin by outlining the theory, starting with a series of propositions underlying the theory. The domain of the theory considered here is obviously single species populations. However, as will also be discussed, the approach used here has been used as well to consider issues related to interacting species.

Propositions

As in other chapters, the development of the theory starts with a series of propositions. Here, these are very general, and in some senses phenomenological. The alternate development below in many ways is more tightly tied to the biology. The propositions here are developed by

considering what ingredients could provide a complete description of single species population dynamics, without specific reference at first to biological processes.

- 1) The dynamics of a population are completely determined if the timing and location and characteristics of all offspring of all newborn individuals as well as all deaths are known.
- 2) The determination of both births and deaths are at an individual level and therefore at the level of an individual birth and death rates can only be given in terms of probabilities. Every birth or death is clearly a chance event and at the most detailed level a deterministic description is never appropriate.

Statement 1) is almost tautological – but informative by pointing out clearly what the simplifications are that need to be made when producing a description that can be effectively used in models. Statement 2) indicates some of the most dramatic problems with developing theories of single species populations. The degree to which averages can be made so that descriptions of births and deaths can be described by simple probability distributions will depend on the importance of various heterogeneities and the size of the underlying population. The description as contained in the propositions above is essentially the basis for a simulation based approach (Grimm and Railsback 2005), which has been variously called individual based modeling or agent based simulation.

Implications and consequences

As will be clear, the unfolding of the assumptions in and simplifications that can be made to these propositions provides a way of understanding much of the subject of single species population dynamics. The role played by proposition 1 will clearly be central.

Specific instances of descriptions of population dynamics arise when particular assumptions are made about the population per proposition , and about the role of stochasticity as embodied in proposition 2. These can really be assumptions about the lack of dependence of population processes on different factors, or can be approximations. One basic approximation that is made in many cases is to describe population sizes using a continuous variable, rather than restricting attention to integer population sizes. If the population size is very large, this approximation is likely to be a good one. Conversely for very small population sizes, descriptions based on birth death models that focus on individual births and deaths will be needed. As another example, one can focus on a mean field (Levin and Pacala 1997) description of a model that includes the distribution of individuals in space, and ignore the role of explicit space in the dynamics of the model. With enough assumptions and approximations, the essentially complete description of population dynamics embodied in these propositions can be simplified to the most basic population models.

For example, assume that all newborns are identical (and even gender can be ignored), generations are discrete and nonoverlapping, the expected number of offspring is the same for every individual and does not depend on the number of individuals, spatial extent is ignored, the environment is constant, and the initial population is large enough so stochastic effects can be ignored. Then, the population description reduces to discrete time exponential (geometric growth). This is the simplest kind of description of population dynamics.

Similarly, one can obtain conditions under which population growth is exponential in continuous time. Here, assume that all processes (death and birth) are independent of age, density, and location (and anything else!), and ignore stochasticity. In this case, the population grows exponentially. Of course, real populations do not behave this way.

It is rather straightforward to see how all simple descriptions of population growth in time and space (i.e., those with age structure, spatial structure, genetic structure, density dependence, or various stochastic influences) fit within this construct as a result of either assumptions that simplify the model, possibly combined with approximations that also simplify the model.. For example, a deterministic model that includes age structure makes implicit assumptions that all individuals contribute to future growth. Also, as a deterministic model with population size as a continuous variable, the model is making the assumption (or approximation) that the population can be well described by a single number rather than a distribution of population sizes. Thus an advantage of this development is that it helps to stress the importance of identifying these approximations and assumptions, which are important to consider when deciding if a particular model applies to a given situation. Continuing with the deterministic age structured population model, one can ask whether the population size is large enough and the environment is constant enough that stochasticity can be ignored. More properly, one would start with a stochastic model, and compute a full distribution of outcomes, and see whether the variance is small enough that the expected outcome is a good description. The development of essentially all single species models thus can be thought of as starting from a microscopic

description, with, of necessity, averaging of various kinds (such as using expected values of a stochastic process) required to produce a macroscopic description that may be of more use in applications or in understanding biological processes.

The development here also makes clear the relationship of population ecology to other subfields of ecology. Proposition 1 provides this insight as well into connections to other parts of ecology. Fields such as behavioral ecology, physiological ecology, and evolutionary ecology all focus on ecological aspects that would affect the reproduction and survival of an individual. Knowing when these aspects at the level of the individual affect the dynamics of populations is clearly a difficult, yet important, question. This idea of the state of the individual is beginning to play a more central role in some models of population dynamics (de Roos et al. 1992), going far beyond the inclusion of age or stage structure. This kind of model structure can include many further descriptions of population state, with the practical constraint that the models developed can become difficult to study or draw conclusions from.

At a higher level of organization, one needs to recognize that populations do not live in isolation. Interactions with other species, or with individuals of different species, clearly affect the dynamics of a focal species through basic processes like competition, predation, parasitism and others. These considerations highlight the issues behind proposition 1. Proposition 1 explicitly excludes the interactions from consideration, while proposition 2 essentially says that these kinds of effects can be lumped into a category called stochasticity. The limitations of the approach described here, where interactions with other species (and physical forcing) are not explicitly included, becomes abundantly clear. Nonetheless, the approach described here can in principle be expanded to multiple species yielding interesting insights, especially in guiding simulation approaches. Yet, the obvious difficulty of using the approach here to develop models of single species, demonstrates the difficulties of building up from this approach to multiple species. Thus community ecology is at a level where alternate approaches may be called for, as described by Holt in Chapter 7.

Example

The stochastic version of the discrete time Ricker model developed in Melbourne and Hastings (2008) provides an illustration of many of the points made here. In that paper, a series of models are developed from first principles. The model is for a population which is discrete in time,

where adults lay eggs, which are subject to cannibalism by the adults. The adults die, and then the remaining eggs which have survived cannibalism are subject to other density independent mortality before the females lay the next generation of eggs.

Of most interest here is the specification of the birth process and other processes in the model, and in particular how various stochastic processes play a role. A series of nested models are developed, where different sources of stochasticity are included, are used, resulting in different descriptions. In particular, the birth (and death) of individuals occurs randomly through time (with a mean rate that is constant per capita), which is known as demographic stochasticity. A second source of stochasticity is the potential variation in rate due to environmental influences, which would produce fluctuations in mean rates between generations. A third source of variability that is less often included is demographic heterogeneity, or differences between individuals in their birth or death rates (due perhaps to different micro-environments, maternal effects, or phenotypes). A fourth source of variability is perhaps the most extreme form of demographic heterogeneity, namely sex differences.

The different combinations of stochastic influences just described lead to a set of eight different models with different combinations of influences. Of most interest here is that these differences at the microscopic level lead to differences in behavior at the macroscopic level. All models have as their mean description the same Ricker model. However, other summary behaviors are different. Mean time to extinction depends critically on the way stochasticity is included, and these results clearly depend on starting from the kind of complete description that is outlined here. With demographic heterogeneity, not surprisingly, time to extinction can be much shorter.

MACROSCOPIC APPROACH

The microscopic approach has advantages in highlighting assumptions, but can be unwieldy and can potentially obscure the relationship to ecological questions. An alternate way of proceeding might be to start with exponential growth and look for those aspects that prevent a population from growing exponentially. Since this highlights different aspects of population biology, I will present this alternative approach in detail as well.

Propositions

Here the propositions essentially start from the simplest description of basic population dynamics.

- 1) A population grows exponentially in the absence of other forces.
- 2) There are forces that can prevent a population from growing exponentially.

This formulation is perhaps less useful than one which takes one step further back, and starts with assumptions that would produce exponential growth as a consequence. Thus, an alternative to proposition 1 would be a series of propositions from which the proposition 1 would be a consequence, combined with the previous proposition 2, now numbered 4.

Propositions (alternative version)

- 1) All individuals in a population are identical.
- 2) The number of offspring (In discrete time) or the birth rate and death rate (in continuous time) per individual are constant through time. In particular these numbers or rates are independent of the number of individuals in the population.
- 3) There is no immigration.
- Real populations are affected by differences among individuals, density dependence and random events.

Although all interesting models build on proposition 4, it is informative to start without this proposition. There are basically two instances of theories obeying propositions 1 through 3, which depend on whether the focus is on discrete time, or continuous time. Either formulation provides an explanation of why proposition 3 is required. In the former case, the population obeys the very simple equation N(t+1) = RN(t), where N(t) is the population size at time t, and R is the per capita number of individuals the following year. In continuous time, the fundamental description is dN/dt = rN, where N is the population size (as a function of time, t) and r is the pre capita growth rate of the population, which is simply defined as the per capita birth rate – the per capita death rate.

Both of these equations have solutions which can be written explicitly, and which take the forms $N(t) = R^t N(0)$ in discrete time, and $N(t) = e^{rt} N(0)$ in continuous time. From these solutions, we can easily see the relationship $R = e^r$.
We note that both of these solutions predict only three qualitatively different kinds of behavior. In particular, if R>1 (r>0) the population grows exponentially, while if R<1 (r<0) the population declines exponentially. The third possibility is the highly unlikely one of perfect balance, R=1 (r=0), where the population would remain unchanged through time. These simple conclusions are very powerful, and form the basis of much of the theory underlying studies of population viability. Yet, this approach is obviously too simple since it ignores many factors that affect populations, and also because it cannot explain (since the perfect balance is highly unlikely) populations that do not grow or decline exponentially.

Under this development, the study of population biology becomes one of understanding the causes of deviations from exponential growth, and then the consequences of the actions of these causes. Thus, this approach in some sense mirrors the approach described in development 1, by building up to the most complex situations, rather than by looking at what simplifications can be made to the complex system to produce a simpler one.

First, I will consider modifications that in some sense are less dramatic in that some aspects of exponential growth are still preserved. In the simplest cases just considered, the population grows exponentially from the beginning and for all time. The next simplest behavior might be for populations to grow exponentially after some initial phase of growth that is not exponential. Or perhaps a quantity other than simply population size grows exponentially.

If individuals in a population are structured by age or stage, but the other propositions all hold (with appropriate modifications to deal with structure), the conclusions of exponential growth can essentially be recovered, under some further assumptions that essentially say that all age or stage classes can be reached by all others at all sufficiently long times (see Caswell, 2001 for a discussion). More specifically, one can find that asymptotically, the population will grow exponentially, but that initially the growth may not be exponential since there may initially be more of an age class that either contributes more or less to immediate numbers in the following year. For example, if all individuals are immature, then it may be several years before a population grows at all. The growth rate is well defined, and can be calculated, in principle for all cases. The idea of three qualitatively different kinds of behavior is essentially recovered, with a population either growing, shrinking or remaining constant. Yet, this seemingly simple and small addition in complexity produces a theory (Caswell 2001) that is so rich and important for

answering questions in population biology that despite a huge literature many questions remain unanswered, although the basic framework is completely understood.

An alternate possibility that essentially retains exponential growth, but that has very different consequences, is to allow for stochasticity (see review in (Lande et al. 2003). First, consider the case of all members of the population as identical, but allow either that the environment varies from year to year, affecting the population growth rate, or that there are demographic stochastic events. Demographic stochasticity occurs when the population is small enough that the chance events of either multiple births in a row or deaths in a row or, in discrete time, variability in offspring number, could have significant impact. In these cases there may still be a quantity that grows exponentially, but it may be uninteresting. For example, it is easy to derive models in which the mean population size grows exponentially, but the probability of extinction is 1. This is the so-called gamblers ruin paradox, which shows up even if the population has a probability of tripling each generation which is equal to the probability of going extinct. Here, the time until extinction is a much more interesting quantity, since this is a more complete description of population behavior. Going further, one could consider the full distribution of population sizes at a given time. This issue is discussed above in the context of the Ricker model.

Before going on to density dependence where per capita birth and death rates depend on the number of individuals, it is intriguing to consider the seemingly simple combination of age structure and stochasticity. Here, even the most basic questions of the growth rate are known only under assumptions that restrict the size of the stochastic effects and the form of stochasticity (Lande et al. 2003), although results do indicate that growth will still be exponential (in the same sense as for stochastic models without structure). Since this is a case that is essential for applications as diverse as conservation biology and fisheries management, the difficulty of analysis highlights the need to be very careful when considering particular instances of theories in population dynamics. Assumptions and approximations used to obtain solutions need to be carefully considered. Including other forms of structure, such as space, or genetic structure would be a next natural step. This too, can be done in ways so that growth is eventually exponential. but for many applications, it will be the transients (Hastings 2004) that are important. Even a model as simple as a description of a population living in two patches, coupled by dispersal where each patch has local population dynamics with overcompensatory density dependence has long transient behavior. Dynamics might be characterized as in phase (population sizes both patches going up and down together) for long periods of time suddenly switching to out of phase dynamics or vice versa, all without any external influence.

It is clear that populations do not continue to grow exponentially. So, the next step is to include factors that prevent exponential growth. Some, such as density dependence, can be considered in the context of single species. Others, such as predation, disease, or competition, require at least implicit, if not explicit, consideration of other species. Even density dependence typically implies that other species or specific resources are of interest. Thus, some of the limitations of single species population dynamics as a self-contained theory come to the fore.

Dynamics depend critically on the form of density dependence. Density dependence that is compensatory would mean that the numbers per capita the following generation are lower, but total numbers still go up as numbers go up, while with overcompensatory density dependence total numbers the following generation go down as numbers in the current generation go up, reflecting overuse of resources. It should be noted that the simple addition of overcompensatory density dependence to population models can lead to wild and complex dynamical behavior (May 1976). (Although this arises in the simplest models only in discrete time, similar behavior is exhibited by more complex continuous time models.) Thus the domain of single species population models is large enough to produce essentially all possible kinds of dynamics.

For many applications, it will be the transients (Hastings 2004) that are important. Even a model as simple as a description of a population living in two patches, coupled by dispersal where each patch has local population dynamics with overcompensatory density dependence has long transient behavior. Dynamics might be characterized as in phase (population sizes both patches going up and down together) for long periods of time suddenly switching to out of phase dynamics or vice versa, all without any external influence.

Eventually, one can build up to the full consideration of all factors included in the microscopic approach. Neither development is primary – it is the connection between microscopic and macroscopic that is central.

METAPOPULATION APPROACH

The theory of metapopulations has continued to play an increasingly important role in the study of single species population dynamics. As it is a single species population approach it should be

possible to place metapopulation models in the context of the approaches developed here. The metapopulation approach is another version of a macroscopic approach that is different from the one just presented. Ideas from the development of structured metapopulation models provide one way to extend the macroscopic description. Although it would be possible to develop a set of propositions specifically for metapopulation dynamics, or add specific ones or revise the sets given above, metapopulation models are a form of single species dynamics and therefore it should be possible to see how metapopulation models are a specific instance of a single population theory.

It is thus informative to think of how to 'force' a description of metapopulation dynamics to fall within the structure of the microscopic approach developed here. One needs to make specific assumptions about the effect of individuals on each other's probability of having offspring and dying as a function of their location in space. Thinking about metapopulations this way also points out an issue with proposition 1 of the microscopic approach. I did not explicitly include a requirement of the knowledge of the state of an individual (e.g., its location) throughout its life. In a metapopulation approach, the location of an individual affects population dynamics through its effect on the production of offspring (both its own and the offspring of others).

As a way of making this discussion more concrete, begin with the simplest metapopulation model, the Levins (1969) model, which takes the form dp/dt = mp(1-p) - ep, where p is the fraction of patches or habitat occupied by the species, m is the per occupied patch colonization rate of empty patches, and e is the extinction rate of patches. It is obviously much easier to see how this model fits into the microscopic approach than the macroscopic one, and there may be more than one way to specify fully a model that looks like the Levins model. One could specify a fixed population level for an occupied patch, and assume that death occurs either by a single individual dying and being replaced from the local patch, or by having all individuals die together. Reproduction would then also be of two kinds. One would be replacing a single individual since with a fixed carrying capacity per patch these are essentially equivalent) that would produce enough individuals to colonize and empty patch and immediately fill it to carrying capacity. The somewhat artificial nature of these assumptions is primarily an indication of the kinds of averaging that go into the Levin's model. In turn, acknowledging and

recognizing these assumptions is an important aspect of drawing conclusions, both for basic understanding and for management, from a Levins model.

A somewhat more straightforward way to connect metapopulation theory with the development in this chapter is to start from models of structured metapopulations (Gyllenberg and Hanski 1997; Hastings 1991; Hastings and Wolin 1989). In this case, the models specify the number of individuals within a patch. Then, the connection to the single species theory becomes clearer. The specification of where offspring of newborns end up is thus a statement about the probability of remaining in the current location, or moving to a new location.

CONCLUSIONS

The study of population dynamics at the level of the single species is an area where understanding can be significantly enhanced by focusing on the fundamental assumptions and conclusions of the theory. In particular, since this is a theory that clearly makes assumptions which may be unrealistic, explicit consideration of the basis for this theory can help make precise domains of applicability. This is particularly important, as many management decisions are based on single species theory, and will likely be based on single species theory into the future.

The two different developments here, the microscopic and the macroscopic, not only provide insights into single species population dynamics, but highlight what are essentially the key issues in much of modeling in ecology. How does one relate events at one level of organization, the population, to those at another level, the individual? How much simplification should be done before writing down a model? What is the appropriate scale to focus on?

All these questions really come to the forefront when using the theory to make predictions in the face of what is often limited data. Stochasticity clearly plays a role. The fact that model complexity is an important issue, and that simpler models may be better for making predictions (and therefore in management) (Ludwig 1999) is often not well enough appreciated. When choosing simpler models, the more formal development of the theory here which highlights the role played by assumptions and approximations can show both the possibilities and limitations of the theory. This is particularly important in management, especially in the context of global change.

The limitations of single species theory are particularly clear when recognizing how much that a single species theory leaves out. The role played by species interactions must be an important one. Yet, management is often still done within a single species context, and insights into bioeconomics (Clark 1990) from single species analyses have been very influential.

Even restricting to the domain of single species theory, the microscopic approach presented here, starting from an essentially complete description of single species population dynamics, exposes the difficulties of developing a complete theory even for this case. A hope is that the formal development is useful in organizing current theory and guiding future developments. Literature Cited

- Caswell, H. 2001, Matrix population models. Sunderland, MA, Sinauer.
- Clark, C. W. 1990. Mathematical Bioeconomics: The Optimal Management of Renewable Resources, Pure and Applied Mathematics series. New York, Wiley.
- Grimm, V., and S. F. Railsback. 2005, Individual-Based Modeling And Ecology, Princeton University Press.
- Gyllenberg, M., and I. Hanski. 1997. Habitat deterioration, habitat destruction, and metapopulation persistence in a heterogeneous landscape. Theoretical Population Biology 52:198-215.
- Hanski, I. 1999, Metapopulation Ecology. Oxford, Oxford University Press.
- Hastings, A. 1991. Structured models of metapopulations. Biological Journal of the Linnean Society 42:57-71.
- 2004. Transients: the key to long-term ecological understanding? Trends in Ecology and Evolution 19:39-45.
- Hastings, A., and C. L. Wolin. 1989. Within patch dynamics in a metapopulation. Ecology 70:1261-1266.
- Kingsland, S. E. 1995, Modeling Nature: Episodes in the History of Population Ecology, University of Chicago Press.
- Lande, R., S. Engen, and B. E. Saether. 2003, Stochastic Population Dynamics in Ecology and Conservation, Oxford University Press, USA.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.
- Levin, S. A., and S. W. Pacala. 1997. Theories of simplification and scaling of spatially distributed processes, Pages 271–296 Spatial Ecology, Princeton University Press.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15:237-240.
- Ludwig, D. 1999. Is It Meaningful to Estimate a Probability of Extinction? Pages 298-310, Ecological Society of America.
- Malthus, T.R. 1798, Principles of Population, J. Johnson, London
- Melbourne, B.A. and Hastings, A. 2008 Extinction risk depends strongly on factors contributing to stochasticity. <u>Nature 454</u>: 100-103

- May, R. M. 1976. Simple mathematical models with very complicated dynamics. Nature 261:459-467.
- de Roos, A. M., O. Diekmann, and J. A. J. Metz. 1992. Studying the dynamics of structured population models: A versatile technique and its application to daphnia 139:123-147.

Chapter 7: Natural enemy-Victim Interactions: Do We Have a Unified Theory, Yet? Robert D. Holt

"Nature red in tooth and claw", "Eat, or be eaten": these and other popular sayings provide everyday testament to the fact that exploitation of one species by another via consumption pervades the natural world. I begin with a few words on the nature of theory in ecology, and then characterize the "domain" of natural enemy-victim ecology. I suggest that any sensible theory of interactions between natural enemies and their victims must be grounded in basic population theory, and that a necessary duality in the foundations of population biology helps identify broad principles for the theory of natural enemy- victim interactions. There is a rich existing theory for this ecological domain but also important unresolved issues and challenges for continued research.

A FEW WORDS ON THEORY IN ECOLOGY

Definitions of "ecology" usually include understanding patterns in the distribution and abundance of organisms, and elucidating the processes generating those patterns (Begon et al. 2006). Scheiner and Willig (Chapter 1) state that a general theory of ecology should contain relatively discrete 'constitutive theories', which ideally include: i) a *domain*, delimiting the range of the natural world encompassed in the theory; ii) a set of *propositions*, or 'foundational concepts' (Berryman 1999, Turchin 2001) that help understand the domain and are related to basic principles in general ecological theory and in other sciences; iii) formal, mathematical *models* that instantiate general ideas, and provide a bridge between these ideas and empirical studies; iv) and finally, *connections* with other 'constitutive theories.' In my view, these desiderata should not be viewed rigidly, but fluidly and flexibly. It may not be possible to adequately describe, understand, or even identify the basic propositions of a theory, without having a formal model at hand. Or, the propositions will themselves emerge from attempts to develop models, an effort that usually requires explicit or implicit linkages with other areas of theory. The boundaries of 'domains' will often be fuzzy. And so on.

The chapter is organized as follows. First, I will reflect on the domain of the body of theory I am concerned with, and touch on prior attempts to articulate a general theory of this

domain. Second, I will briefly discuss two foundational concepts from the general theory of ecology, each of which can be serve as a basis for general natural enemy-victim theory, and both of which in my view are necessary for a well-rounded theory of this domain. Third, I will develop a series of propositions, made more crisp by being cast in the language of a very general, formal model. I will present propositions which I suggest are universal for natural enemies specialized to a single kind of victim, and should thus be contained within any coherent specific model of a natural enemy-victim interaction, and then turn to a number of other propositions which are broad generalizations, but not without exceptions. Fourth, I will discuss some important issues that have arisen within key subdomains of the general theory, namely predator-prey and host-pathogen theory. Some key issues are surprisingly still unresolved, even after much debate. Fifth, I will discuss how some of the propositions of the theory need to be qualified when considering generalist natural enemies. I conclude with pointers as to future directions natural-enemy victim theory is likely to take.

THE DOMAIN OF THE THEORY OF NATURAL ENEMY-VICTIM INTERACTIONS

What is the "domain" of natural-enemy victim theory? Murdoch et al. (2003, p. 1) remark "The consumer-resource interaction is arguably the fundamental unit of ecological communities. Virtually every species is part of a resource-consumer interaction, as a consumer of living resources, as a resource for other species, or both." This is not quite right. Resource-consumer theory legitimately includes plants as consumers of light and nutrients, and decomposers as consumers of dead organic matter. The R*-rule (Tilman 1982) is a proposition of general consumer-resource theory, for instance for phytoplankters competing for nutrients. Resource-consumer theory has intimate ties with other areas of ecological theory such as foraging theory, competition, and niche theory, all of which focus on the consumer.

I suggest that the "domain" of natural enemy-victim theory spans all resource-consumer interactions where the resource is alive, and is typically harmed by that consumption. The focus is as much on the resource (victim) as on the consumer (its enemy). This domain includes multiple "subdomains". Much of modern ecology has grown out of pragmatic pursuits – for instance, harvesting populations for meat or fur, or the mitigation of epidemics, or the control of agricultural pests. Moreover, the classical description of ecosystems in terms of primary producers, primary consumers, and secondary consumers focus on major biological differences

among species in the kinds of resources they require, leading to natural suture zones in theory. Because of this, there are many *particular* subtheories that have arisen dealing with the exploitation of different kinds of 'victims' by different kinds of 'natural enemies'. Yet there are many common features of these particular theories. General natural enemy-vicitm theory tries to identify these common unifying threads, and also to provide a conceptual structure for understanding and organizing the differences among these more particular theories. For reasons not entirely clear to me, the term 'natural enemy' is more frequently encountered in the United Kingdom than in the United States (e.g., Crawley 1992). The phrases 'natural enemy – victim' or 'exploiter -victim' provide succinct umbrellas encompassing several distinct arenas of ecological theory, including: "true" predator-prey theory (Taylor 1984, Turchin 2003), hostparasitoid theory (Hassell 1978, 2000, Murdoch et al. 2003), host-pathogen (= microparasite) and host-parasite (=macroparasite) interactions (Hudson et al. 2002, Keeling and Rohani 2007; including infectious disease epidemiology, e.g., Anderson and May 1991)), and plant-herbivore interactions (Crawley 1983, Owen Smith 2002). The distinction among these bodies of theory reflects the functional details of how the exploiter engages its victim, and its impact upon victim fitness. One important question that all of these subdomains grapple with is "static": what is the average effect of each species in limiting the abundance of the other species? Another broad set of questions has to do with the "dynamics" generated by the interaction, i.e., is the interaction stable? if not, do unstable populations tend to cycle or show more complex behaviors?

There is an enormous, increasingly mature, and conceptually complex body of ecological theory related to natural enemy-victim interactions. There are lively debates not just about fine points, but about major and indeed central features of these interactions. This chapter can at best point to a few general principles and unifying structural features of this sprawling theoretical edifice. Much of what I will say attempts to synthesize and reiterate perspectives ably presented at various places in the ecological literature (e.g., Crawley 1982). How can one pretend to synthesize across this vast arena, and do so in a way that does not sound like a regurgitated abstract of an encyclopedia? In empirical ecology, there is a robust discipline of meta-analysis that assesses patterns and hypotheses by bringing together results across disparate experimental studies. There is not as yet a comparable rigorous 'metatheory' for gleaning (using formal protocols) conclusions from a range of specific models. In the absence of such a formalized

approach, there are two broad approaches to synthesizing complex realms of ecological theory, both of which I think are essential steps towards general theory.

First, one can take the stance championed by Schoener (1986) towards community ecology and identify major axes of variation distinguishing different kinds of natural enemyvictim interactions. May and Watts (1992) for instance provide a table comparing life histories and other biological attributes across major categories of natural enemies. Lafferty and Kuris (2002) suggest four axes for characterizing parasites and predators, such as: how many victims are attacked in the enemy's lifetime; does the enemy reduce victim fitness to zero (e.g., true predators, parasitic castrators)? Such salient differences should be reflected in the structure of models describing their dynamics (Hall et al. 2008). Space precludes my following this particular direction of theory synthesis.

Second, one can attempt to identify general cross-cutting themes by comparing a range of more particular models (as in e.g. Murdoch et al. 2003, Holt and Hochberg 1998, Borer et al. 2007), highlighting overarching constraints within which all such models must operate. This relies on the instincts of a practitioner, rather than any kind of formal comparison. It can be helpful to start out by returning to fundamentals, to clarify the assumptions and flow of logic that define the lineaments of theory. This is what I attempt here. Murdoch et al. (2003) conclude their fine volume by saying "Consumer-resource theory [which they use broadly congruent with what I mean by "natural enemy-victim theory"] is... quite unified and internally coherent." I think they are right ... in part, and with caveats. Not all subdomains of natural-victim theory are welldeveloped theoretically, and potentially significant qualitative differences exist among subdomains that may complicate unification. Even existing theory for "true" predator-prey and host-parasitoid interactions (doubtless the most fully developed of the lot) has key gaps, and there are basic disagreements about some important issues.

Prior syntheses of natural enemy-victim theory

Several authors have attempted to develop fundamental properties of subdomains of naturalenemy victim theory. For instance, Murdoch et al. (2003, p. 76) list "fundamental properties" they believe characterize predator-prey theory based on continuous-time models:

1. Interacting predator and prey populations tend to oscillate

2. Satiation in predators, and time lags in predator response, increase the tendency to oscillate;

3. Direct negative density in vital rates tends to dampen oscillations;

4. Combination of factors in #2 and direct density dependence in prey tends to produce limit cycles

5. Most processes suppressing oscillations also increase prey equilibrium density.6. Spatial heterogeneity and/or age structure can lead, via asynchronous trajectories of component populations, to indirect density dependence which may stabilize populations.

This list is an important statement of potential generalities about this significant subdomain of natural enemy-victim theory. Yet they are not really foundational propositions, but rather emergent conclusions, broad generalizations about predator-prey interactions. To some extent, as we will see, these generalizations may extend well beyond continuous-time predator-prey models to other classes of natural enemy-victim interactions. Moreover, one can find counter-examples or limitations in these generalizations. For instance, many (possibly most) predator and prey populations do not oscillate; predation satiation can simply prevent predators from regulating prey at all, rather than causing cycles; certain kinds of time lags in predators can be stabilizing (Nunney 1985); spatial heterogeneity can likewise at times destabilize (Hochberg and Lawton 1990; Holt 2002); and in discrete generation models, strong negative density dependence creates its own form of instability (May and Watts 1992). The above statements are not thus universal laws, but more like central tendencies in a field of theoretical possibilities. Below, I will attempt to identify a few general propositions that arguably should hold for all, or at least very many, natural enemy-victim systems.

FOUNDATIONAL CONCEPTS

To see our way forward to identifying both elements of conceptual unification, and of discord, it useful to go back to basics. All natural enemy-victim theory necessarily rests upon the foundation of population dynamic theory (Hastings, chapter 6). There are two complementary approaches to describing population processes, reflecting an essential duality in the nature of life itself (Eldredge 1985). The first half of the duality emphasizes information; the second, physical processes. Both are needed elements in the foundation for a comprehensive, even partially unified, theory of natural enemy-victim interactions.

Demography and exploitation

Life involves the replication of organisms, which are transient physical vehicles transmitting packets of heritable information through time (Dawkins 1982). Natural enemy-victim theory is an elaboration of population dynamic theory that makes explicit the reciprocal dependence of births and deaths in each species upon the abundance of the other species. An explicit consideration of demography is thus essential to providing a firm theoretical foundation for natural enemy-victim theory. As discussed in detail by Hastings (chapter 6), the basic equation of demography is that the numbers of individuals N of a given 'type' in a particular place always changes over a small period of time according to some variant of:

 $\Delta N = B - D + I - E = births - deaths + immigration - emigration$. Because births and deaths happen to individuals, total births and deaths are best expressed as per capita rates. In continuous time in a closed population, where all individuals have equivalent traits and there are enough of them to ignore demographic stochasticity, we have

dN/dt = N(b-d) = Nr, (1)

where *b* and *d* are instantaneous per capita birth and death rates, and *r* is net per capita growth rate. If *r* is constant (i.e., the environment is unchanging), we have the foundational demographic principle of exponential growth (see Hastings, chapter 6): r<0 implies extinction, whereas r>0 leads to unbounded growth. The exponential model is the basic model on which the rest of population biology hangs. Much of the theoretical complexity of modern demography arises from dealing with internal structural heterogeneities in populations; likewise, incorporating the details of population structure is an important objective of much current work in natural enemy-victim theory (e.g., Murdoch et al. 2003; de Roos et al. 2003). For a closed population to persist, and neither go extinct, nor blow up to infinity, on average its births must match its deaths (Royama 1992) -- whether or not its dynamics involve regulation to a tight equilibrium, or cycles, or chaos, or bounded stochastic wanderings. This observation is the starting point of analyses of many important features of natural enemy-victim interactions.

The physical basis of life

Life is a physical process. Organisms are physico-chemical machines, out of equilibrium with their environments, dependent upon recurrent supplies of energy and materials for survival and reproduction. Because natural enemy-victim interactions by definition involve forced transfers of energy and materials from one organism to another, a biophysical perspective is an essential complement to purely demographic models. Both are necessary ingredients in any unified theory of natural enemy-victim interactions.

An important (if minority) tradition in population dynamics emphasizes as a foundational principle the material basis of life as a set of resource transformations, rather than demography (key references include Lotka 1926, Getz 1984, 1993, Ginzburg 1998, Gutierrez 1992, Kooijman 1993, Owen-Smith 2002, Sterner and Elser 2002, Ernest et al., 2003, Brown et al. 2004). Owen-Smith (2005) for instance remarks, "From a biomass perspective, population growth is not the result of a difference between births and deaths (despite the appearance of this statement in most textbooks), but rather of the difference between rates of uptake and conversion of resources into biomass, and losses of biomass to metabolism and mortality (Ginzburg 1998). The biological law of regeneration ...underlies exponential population growth." The second law of thermodynamics implies a continual degradation of living tissues, mandating maintenance and repair. Because this occurs with error, mortality is eventually inevitable, and of course mortality arises from many other factors. Among the most predominant of these factors are natural enemy-victim interactions.

The physical processes of life mandate the extraction of resources from the environment, and resource availability varies across space and through time, so it follows that birth and death rates should also vary with environmental circumstances. Some organisms (producers) meet their needs entirely from the nonliving environment. But, the presence of other organisms provides a tempting packet of materials and energy, packets that can be exploited. Because of the conservation of mass and energy, such exploitation amounts to an at best zero-sum game – which lies at the heart of natural-enemy victim interactions, and can be stated as the first proposition of the theory:

Proposition 1: Material resources acquired by consumption that benefit one organism (the natural enemy, or exploiter), automatically reduce the resources available to another organism (the victim), from which the resource has been taken.

This sets up an automatic conflict of interest, driving behavioral struggles, population dynamics, and coevolutionary races. Much of the splendid diversity of life revolves around the manifold ways such cross-organismal exploitation and struggles play out, and the literature of

natural history and ecology is replete with tales of the details of exploitative strategies and defensive counter-moves (e.g., Caro 2005).

If we consider "true" predators attacking prey, the primary, first-order effect is to limit prey growth by acting as a mortality source; more generally, natural enemies tend to reduce the growth rate of their victims. In the next few paragraphs, I follow a traditional demographic approach to natural enemy-victim interactions and use the physical perspective to identify general constraints that should pertain to any reasonable model. These constraints can be cast as propositions.

PROPOSITIONS AND FORMAL MODELS IN NATURAL ENEMY-VICTIM THEORY

Exploitation has demographic consequences for both the natural enemy and its victim. It is useful to stand back from all the details, and attempt to identify general propositions that pertain to all exploitation systems. The formulation of such propositions is clearest when cast in the formal language of mathematical models. The schemata of population growth of model (1) generalizes naturally to natural enemy-victim systems. We consider a single natural enemy species, attacking a single victim species. In classical predator-prey theory (stemming back to A.J. Lotka 1926 and Vito Volterra (1926, 1931; see e.g., Rosenzweig and MacArthur 1963), models are cast in continuous time, with continuously overlapping generations for both the predator and its prey. In textbooks, it is traditional at this point to introduce the Lotka-Volterra model, which one then makes more complex in various ways. Instead, I will try to work with a general model, for a bit, and then turn to the classical approach. For a natural enemy (e.g. predator) of density P, interacting with a victim of density N, in environment E, the simplest general continuous-time model in a closed system has the form:

$$\frac{dP}{dt} = Pf_P(N, P, E)$$

$$\frac{dN}{dt} = Nf_N(N, P, E)$$
(2)

Here, f_P is the per capita growth rate of the natural enemy, and f_N is the per capita growth rate of the victim. The "environment" includes all abiotic and biotic factors (including other species) that can be assumed to be fixed parameters, or embedded in the functional form of the equations, rather than variables. Model (2) expresses one of the propositions of Scheiner and Willig (2008), which is that the strength of interactions should depend upon the environment. We will mainly

assume the natural enemy is a specialist, but towards the end will also make a few remarks about trophic generalists. (Below, for the sake of variety I sometimes use "predator" and "prey" for "natural enemy" and "victim"; the claims made are mean to apply to say herbivores on plants, as much as "true" predators and prey such as wolves attacking moose.)

There are numerous assumptions at play in models such as (2), matching the assumptions of simple exponential growth models (see Hastings, chapter 6). In each player, it is assumed that one can ignore variation among individuals (e.g., stage structure, genetic variation) and spatial location (a 'mean-field' assumption, so each population is well-mixed); that abundances are sufficiently large to neglect demographic stochasticity; and that reproduction and mortality are continuous processes (rather than say seasonally pulsed). Growth rates are assumed immediately responsive to current densities, with no lags due to development or accumulated metabolic stores.

Despite this somewhat worrisome list of assumptions, just like exponential growth, analysis of the family of models given by (2) provides an essential springboard. The model could in principle apply to many different kinds of natural enemy-victim interactions. For instance, "*P*" could denote a herbivore, and "*N*" the plant it is attacking; or, "*P*" could refer to infected individuals, and "*N*", susceptible individuals within a host population carrying an infectious disease. For the most part, I focus on "classical" predator-prey interactions, but at the end turn to infectious diseases to highlight commonalities as well as differences among subtheories that may hamper completely unified theory of natural enemy-victim systems.

An important caveat

It should be stressed that starting with a pair of coupled differential equations such as (2) makes the assumption that the important time lag in the system arises because of delays in feedback via the interaction. Other model formulations have feedbacks that arise because of discrete time, delays in direct density dependence, and delays that emerge from demography (e.g., discrete juvenile periods). Another way one could attempt to formalize natural enemy-victim theory is in terms of the assumptions made about the nature of time lags (I thank a reviewer for this observation), and this is in effect what Murdoch et al. (2003) have done in their monograph. One of the basic points discussed below is that natural enemy-victim systems are expected on theoretical grounds to oscillate. However, they way in which this happens is influenced in detail-dependent ways on how populations are structured, including both developmental and regulatory delays, and internal heterogeneities such as spatial structure and genetic variation. Developmental delays in consumers mediated by resource consumption can lead to short-amplitude cycles in natural-enemy victim systems, for reasons quite different than those emphasized in classical theory (McCauley et al. 2008).

The complexities of modeling natural enemy-victim theory arise in large part from the expressions one uses to express births and deaths in each species as a joint function of each others' abundances of species interaction, and the environment. It is useful to see how far we can go, without specifying more detail than is absolutely necessary.

Fundamental constraint theory: the "statics" of natural enemy-victim interactions

One rich vein of complexity in modeling natural enemy-victim theory arises from from the expressions one uses to express births and deaths in each species as a joint function of each others' abundances (i.e., their interaction) and the environment. It is useful to see how far we can go, without specifying more detail than is absolutely necessary. Almost by definition of the domain of this theory, the natural enemy should benefit from the victim, whereas the victim suffers reduced growth because of the natural enemy. In equations, were one to monitor each species after reciprocal removal experiments, over the short-run one expects that (for simplicity, we suppress *E*):

$$f_{P}(N,P) > f_{P}(0,P),$$

 $f_{N}(N,P) < f_{N}(N,0).$
(3)

The interspecific relationship thus has a (+,-) sign structure.

If the system is in demographic equilibrium (denoted by an asterisk), $f_P^* = f_N^* = 0$, hence $f_N(N^*,0) > 0 > f_P(0,P^*)$. In other words, at equilibrium, when one removes the predator (leaving everything else fixed) the prey increases, conversely, were we to remove the prey, the predator on its own decreases. The demographic balance is maintained by the (+,-) nature of the interaction. Though (3) is quite reasonable, it may not always hold for generalist natural enemies (see Holt (1983) for more details, and examples). If generalist predators have a suboptimal (= maladaptive) diet, removing one or more prey types may actually (at times) boost predator numbers. A series of fundamental constraints on natural enemy-victim interactions arise from considering the joint implications of the demographic and biophysical foundations of life. Once stated, these seem obvious, but one role for theory is to lay bare propositions that everyone can agree upon, which permits us to focus on issues that provide grist for the mill of continued disputation.

Proposition 2. No species can rely entirely upon self-consumption, and persist. This follows from the second law of thermodynamics, which states that useful energy is converted to useless energy (heat) whenever work is performed, coupled with the demographic constraints required for persistence. Reproduction, simply staying alive, and consumption itself -- all are varieties of work. With self-consumption, the amount of biomass created must necessarily be less than the biomass consumed. So, if we start with a certain number of individuals in a population, and they only have themselves to consume, after each act of consumption there will be a net loss of biomass. Over time, a population feeding only on itself will inexorably decline to extinction. So, one will never expect to see a community comprised entirely of cannibals – at least not for very long! This principle is the ecological analogue of the impossibility of a perpetual motion machine. I am not claiming that it is impossible for such an organism to exist at all. A modern Dr. Frankenstein in her genetic engineering laboratory could maybe make such a population and get it off the ground– but it could not persist. Indeed, there are some infamous episodes in our own species' history (e.g., the Donner party of American emigrants bound for California, who were caught in an October snowstorm in the Sierra Nevada, ran out of food, and resorted to cannibalism to survive; Diamond 1992) where we created a transient one species natural enemy-victim interaction.

Proposition 3. *There is a necessary asymmetry in the "benefit" (measured in terms of energy and materials) gained by the exploiter, and the "cost" (in the same units) inflicted on the victim; in particular, the "cost" always is greater in magnitude than the "benefit"*. This may not imply an asymmetry as measured by Darwinian fitness, which is measured as descendents per organism, rather than per unit of material. (This is why I put the terms "cost" and "benefit" into quotes above.) The part of this statement that refers to energy again follows from basic thermodynamics. I am not aware of a formal proof of the part of this statement that deals with material pools (e.g., nitrogen content in body tissues), but it would appear to be a very robust empirical generalization. From conservation of mass, the benefit gained by the exploiter cannot

exceed the cost imposed on the victim. Moreover, there is always some wastage in consumption, so the gain in, say, nitrogen for a consumer must always be less than the loss experienced by the consumed.

In model (2), if *P* and *N* are measured in equivalent units (e.g., energy content in each population), then this proposition implies that

$$P(f_{P}(P,N) - f_{P}(P,0)) < |N(f_{N}(P,N) - f_{N}(0,N))|.$$
(4)

Proposition 4. On average, the energy and nutrients gained from each act of consumption must exceed the amount of energy and nutrients used by the natural enemy in that act, if the consuming species is to persist. Proposition 3 is a statement about the aggregate consequences of exploitation across trophic levels. Proposition 4 is superficially similar, but subtly different; it applies to average acts of consumption, within the consumer level itself. Before any 'act' of consumption, an individual natural enemy starts with a certain amount of 'capital' energy and nutrient reserves within its body. Some of this capital is expended during consumption – after all, consumption and assimilation into body tissues are both forms of work, and so operate within the constraint of thermodynamics, which implies a fundamental rock-bottom inefficiency. If these expenditures are not matched by gains from consumption, there will be a loss in biomass with each act of consumption, and if this loss continues, act after act, extinction is inevitable. Proposition #3 is a necessary, but not sufficient, statement about consumer persistence; the net amount of energy and nutrients gained per act of consumption must not only be positive, but be sufficiently great to exceed other costs (e.g., inherent mortality, losses to predation or disturbance, etc.). I will return to Proposition 4 briefly below.

Proposition 5. *The immediate effect of each act of consumption is to reduce the biomass of the consumed victim.* This proposition is almost a restatement of part of the definition of consumption of a resource, yet it has important consequences that are worth focusing on explicitly. In particular, with continued consumption, for victims initially at equilibrium, without compensatory feedback there is a decline towards extinction. If we consider a victim population initially in equilibrium (so births match deaths), and add recurrent consumption by a natural enemy, with no compensatory changes such as additional births or regrowth, or declines in consumption, the victim will inexorably go extinct. Characterizing compensatory feedbacks that affect victim species is thus a vital part of understanding the dynamics of natural enemy-victim interactions.

General (but not universal) dynamical propositions

To go beyond these schematic assertions, and begin to make mathematical models that can be related in more detail to concrete empirical systems, more assumptions are needed. The next two propositions are generally assumed in natural enemy-victim theory, but do not actually necessarily hold.

Proposition 6. *An increase in the victim population increases the rate of consumption by each individual natural enemy*. In the subdomain of predator-prey theory, the functional response of an individual predator to its prey is the rate of consumption of the prey, as a function of prey abundance (Holling 1959), so proposition 6 would state that the functional response can be described as an increasing function of prey density.

Proposition 7. *The increased consumption generated by increased victim abundance in turn fuels an increase in the per capita growth rate (fitness) of the natural enemy population.* In predator-prey theory, the "numerical response" of the predator describes how its own population size increases with prey density, so proposition 7 states that the numerical response is an increasing function of prey abundance. More specifically, the numerical response of the predator should be expressed as a function of its functional response (the "biomass conversion" principle proposed by Ginzburg 1998). If Proposition 7 holds, in model 2,

$$\frac{\partial f_P}{\partial N} > 0. \quad (5)$$

Going back to the basic models of Lotka (1926) and Volterra (1926, 1931) (in continuous time), and Nicholson and Bailey (1935) (in discrete time), inequality (5) has been a basic assumption of natural enemy-victim theory for a very long time. Inequality (5) is not a universal "law" of trophic ecology, but it is certainly a plausible generalization that applies to many natural systems. In the absence of specific knowledge to the contrary, Proposition 5 is a reasonable starting point for modeling.

But it is useful to note significant exceptions to (5). One comes from studies of large grazing herbivores. Energetic intake rates can begin to decline at sufficiently high grass biomass because of foraging constraints, and a correlation between forage digestibility and biomass (Wilmshurst et al. 1999, 2000). Other exceptions arise when prey can engage in group defenses, which may be more effective at high prey densities (Caro 2005). Finally, an important class of

exceptions comes from host-pathogen systems involving vectors, where an increase in host numbers, given a relatively fixed number of vectors, can dilute the likelihood of pathogen transmission to a host (Ostfeld and Keesing 2000), leading to a nonmonotonic relationship between host population size and disease transmission; there may even be host population sizes sufficiently large for a pathogen to die out because the dilution of vector attacks provides insufficient transmission between infected and susceptible hosts (Randolph et al. 2002).

Proposition 8. All else being equal, the total rate of consumption of the victim (and hence the magnitude of its loss due to consumption) increases with the abundance of the natural enemy. In terms of model (2), this means that

$$\frac{\partial f_N}{\partial P} < 0 \quad (6)$$

Again, this is built into nearly all models built on the Lotka-Volterra framework. And again, this inequality is a reasonable generalization, but it is not a universal law. There may be exceptions; for instance at high predator numbers there may be strong direct interference among predators, so that the total mortality imposed upon prey actually declines with increasing predator numbers. As a limiting case, inequality (6) may not strictly hold with "pure" ratio-dependent predation (see below). Victim behavioral responses can sometimes imply that the victim suffers fewer losses, at higher densitites of the natural enemies (Abrams 1993), and that when both species have adaptive behaviors, each can at times benefit the others (Abrams 1992).

Proposition 9. In the absence of the focal victim, a specialist natural enemy goes extinct (in a closed environment). We noted above that at equilibrium in (2), were we to carry out removal experiments, we would observe that $f_N(N^*, 0) > 0 > f_P(0, P^*)$. Because we have assumed that the natural enemy is a specialist, given that reproduction requires energy and materials, there will be no births in the absence of the required victim species, and because all organisms are mortal, there will still be deaths, hence the right hand inequality holds, for all *P*. This implies the natural enemy goes extinct.

Proposition 10. *There is context dependence in natural enemy-victim interactions*. The quantitative relationship between the aggregate rate of consumption of the victim, and natural enemy density, and between victim abundance and consumption, will nearly always vary with environmental context, and with the presence and abundance of other species. This proposition differs from Propositions 1-5, because it does not follow from more basic law-like propositions

in demography and physical biology. It is like Propositions 6 and 7, because it is an inductive generalization across many empirical studies (and matches a proposition in Scheiner and Willig (2008)). It differs from 6 and 7, because it is schematic, and without further fleshing out, it cannot be used to inform dynamical theory, other than to caution that "different things will happen in different circumstances." However, stating this as a formal proposition is useful, as it alerts researchers to the importance of identifying the major dimensions of potential context dependence in their empirical studies.

Proposition 11. Natural enemy-victim interactions have a propensity towards unstable dynamics if the victim is strongly limited by the interaction and the natural enemy is a specialist. Given that the predator limits prey growth, and that the prey boosts predator growth rate, a second key issue revolves around the ability of the predator to regulate that prey (i.e., keep prey numbers within bounds), and of the prey in turn to regulate the predator. When the natural enemy is ineffective at limiting its victim, this is not an issue: the regulation of the system is determined by bottom-up forces or by factors intrinsic to the prey. If, by contrast, a natural enemy effectively limits its prey to low levels, a generic property of simple models of these interactions is that they tend to oscillate, because of delayed density dependence – an increase say in predator numbers will eventually depress predator growth rates, but only after the prey are depressed in abundance, which takes some time; in like manner, an increase in prey numbers leads to an eventual increase in prey mortality, but with a lag dependent upon the time scale of the predator's numerical response. Turchin (2001) has suggested that one of the general laws of ecology is that *population cycles are likely when effective consumers exploit living resources*. Murdoch et al. (2003, p. 70) likewise list a tendency towards oscillatory behavior as first in their list of "fundamental properties" of predator-prey models. I think these authors are right, and that the propensity for oscillatory dynamics should be viewed as a basic principle of natural enemyvictim interactions. (And do note the cautionary remarks above .)

Such oscillations can arise at many temporal scales. An individual prey for instance may have behavioral responses to predators, and the predator may likewise response. Such behavioral responses can set up strongly unstable spatial dynamics (Abrams 2007). The simplest models of predator-prey and host-parasitoid population interactions tend to oscillate (see below). Seger (1988, 1992) has argued that the evolutionary dynamics of asymmetric antagonistic interactions, such as between a predator and its prey, or between a host and its parasite, are inherently unstable, leading to a constant coevolutionary churning in gene frequencies and traits. This is particularly the case if fitness in each species is determined by one to a few major loci and are interspecifically frequency-dependent (e.g., the fitness of defense morphs in a prey species may vary with the frequency of attack morphs in a predator).

Whether or not oscillations damp out following a perturbation, or instead settle into stable cycles or more complex patterns of dynamical behavior (e.g., chaos), depends upon many system-specific details. To see why there is a generic tendency towards oscillations, we start with a system in equilibrium and examine how it responds to small perturbations in abundances (and here I follow well-trod ground that has been covered often in the literature, e.g., Murdoch and Oaten 1975, Murdoch et al. 2003). By linearizing around the equilibrium, the local stability of the system given by model (2) is determined from the eigenvalues of the Jacobian matrix *J* (where the partial derivatives are evaluated at the equilibrium):

$$J = \begin{bmatrix} N * \frac{\partial f_N}{\partial N} & N * \frac{\partial f_N}{\partial P} \\ P * \frac{\partial f_P}{\partial N} & P * \frac{\partial f_P}{\partial P} \end{bmatrix} = \begin{bmatrix} ? & - \\ + & ? \end{bmatrix}$$
(7)

What the left-side matrix does is compactly describe the impact of small perturbations in abundance on the growth rate; the diagonal terms describe the impact of each species upon itself (viz., direct density dependence), and the off-diagonal terms, the effect of each species on the other, (cross-species density dependence, which provide one measure measure of interaction strength). The properties of this matrix provide a linear approximation of the full model (2) (which may be highly nonlinear) near equilibrium, and in particular determine whether small perturbations dampen away, or instead grow, and also if the dynamics near equilibrium show oscillations, or not. The matrix on the right of the equal sign describes the sign structure of interactions for this species pair. The signs of the off-diagonal terms follow from Propositions (6) and (5); as we will shortly see, the signs of the diagonal terms may also emerge from the natural enemy-victim interaction, as well as from extrinsic forces acting on the populations.

Following well-known recipes (e.g., Otto and Day 2007), if the trace of J,

$$T = N * \frac{\partial f_N}{\partial N} + P * \frac{\partial f_P}{\partial P}, \quad (8)$$

is negative, and the determinant D is positive, so that

$$\frac{\partial f_N}{\partial N} \frac{\partial f_P}{\partial P} > \frac{\partial f_N}{\partial P} \frac{\partial f_P}{\partial N}, \quad (9)$$

then the real part of the dominant eigenvalue of J is negative, and the system is locally stable. The quantity T is a measure of the overall strength of direct density dependence added up over both species, weighted by their equilibrial abundances. Quantity (9) in a sense states that direct density dependence is stronger than the strength of the interspecific interactions. Thus, stability hinges on the nature of direct density dependence in one, or both, species. If both predator and prey experience direct, negative density dependence, both (8) and (9) can hold, and the interaction is locally stable. If neither holds, then it is unstable. If one holds, and the other does not, then stability is conditional. We will tease out a bit more of this conditionality in a few paragraphs.

Whether or not oscillations occur is determined by the sign of the quantity

$$Q = \left(N * \frac{\partial f_N}{\partial N} - P * \frac{\partial f_P}{\partial P}\right)^2 + 4N * P * \frac{\partial f_N}{\partial P} \frac{\partial f_P}{\partial N}.$$
 (10)

If Q < 0, oscillations will arise following a perturbation (the dominant eigenvalue of (*J*) has complex parts). If neither the natural enemy nor the victim has a direct effect upon itself (i.e., direct density dependence is negligible), the diagonal terms are zero, implying that T=0. From Propositions 7 and 8, the remaining term is negative, so Q < 0, and oscillations in abundance will surely occur following perturbation. In this limiting case, the oscillations have no tendency to return to the equilibrium. If at a stable equilibrium density dependence is relatively weak, Q will still be negative, and damped oscillations will then arise following a perturbation.

Thus, a generic property of natural enemy-victim systems is that they harbor a tendency to exhibit oscillatory behaviors following perturbation because of the (+,-) structure of the interaction matrix. Kolmogorov (1936; see May 1973, pp. 86-92) elegantly went beyond the linearized analysis presented above to show that given a broad set of reasonable conditions (given on pp. 87-88 in May 1973; these include Propositions 7 and 8 above), models of the form given by (2) above settle into either a stable point equilibrium, or a stable limit cycle. More realistic models that include time-lags in predator responses to their prey (e.g., due to developmental lags) often make such oscillations more likely (Murdoch et al. 2003). Specialist natural enemy-victim interactions can thus be strongly destabilizing. Whether or not these oscillations dampen out, or not, depends upon the strength of direct density dependence in each

species, and whether or not this density dependence is sufficiently negative at equilibrium. In some circumstances (as elaborated below), such density dependence emerges from processes withing the domain of natural enemy-victim theory itself; in others, the density dependence needed for stability is imposed, as it were, from outside the domain.

For bioenergetic reasons it is often the case in predator-prey interactions that $P^* < N^*$, unless prey are highly productive and predators have much longer generations times. Empirical studies of vertebrate predator and prey abundance show that typically (again, not universally) for a given body size, predators are rarer than prey (Marquet 2002). When measured in equivalent units (biomass), mammalian predators are on average two orders of magnitude rarer than prey (Carbone and Gittleman 2002), and since mammalian predators tend to be larger than their prey, the difference in abundance is even greater. By inspection of *T* and *Q*, these observations suggest that density dependence in the prey should often be much more important in determining stability of the trophic interaction, than is density dependence in the predator. However, in other natural enemy-victim interactions, such as parasitoid attacking hosts, or caterpillars chomping on trees, this inequality will be more nearly an equality, or even be reversed.

As a cautionary note, Proposition 11 does not imply that when one sees a natural enemyvictim cycling, the delayed feedbacks embodied in (7) are the reason. As noted above, the cycles might instead be due to intrinsic processes in the victim or natural enemy (e.g., single-generation cycles due to stage structure (Murdoch et al. 2003)), or the victim's interaction with its own resources, or other factors such as maternal effects (Inchausti and Ginzburg 2009).

STEPS TOWARD SPECIFICITY

This is about as much as we can squeeze out of the general model provided by (2). We could just stop here, but this would be a bit of a cheat, since some of the most contentious and interesting issues in natural enemy-victim theory arise when one adds specificity to this general model. The goal of more detailed natural enemy-victim theory is to characterize how specific mechanisms of interactions translate into average effects of each species on the other, as well as positive and negative density dependence within and across species, for each interacting species. Much of the development of specific theory in this area has been guided by the desire to develop a systematic theory for understanding how effective specialist natural enemies manage to coexist with their victims, given the inherent instability that seems to lurk at the heart of these interactions. In addition to its importance for addressing basic ecological questions, this issue is at the heart of devising effective and sustainable biological control programs.

Sources of direct density dependence in natural enemy-victim systems

Before considering specific models, one can broadly divide the sources of direct density dependence (the diagonal terms of (7)) into two classes. On the one hand, density dependence may occur for reasons that are *extrinsic* to the predator-prey interaction, itself. For instance, prey may compete for limiting resources, or predators may experience direct aggressive interference. Such mechanisms of direct density dependence can lead to stability of the joint trophic interaction. Alternatively, there may be positive density dependence in either species, due to Allee effects (e.g., the need to find mates in sexual species). This is destabilizing. On the other hand, direct density dependence in both species may arise for reasons that are *intrinsic* to the predator-prey interaction.

The ingredients needed for analysis may be a challenging list

Crawley (1992) suggests that a minimal list of desiderata for understanding a specialist predatorprey interaction would include: 1. The intrinsic rate of increase of the prey; 2. The functional response of the predator; 3. The predator's spatial foraging behavior; 4. The nature of prey density-dependence; 5. The nature of predator density-dependence; 6. The densities at which the predator and prey isoclines intersect; 7. The relative slopes of the two isoclines at the intersection; 8. The rate of predator immigration; 9. The size of the prey refuge; 10. The rate of between-patch prey dispersal. There are few empirical systems where we have all this information. Given the importance of this topic, this is a rather sobering statement about the adequacy of the knowledge base of our field.

A hierarchical family of Lotka-Volterra models for the predator-prey subdomain

To become more concrete, we return to the traditional starting point for predator-prey interactions in spatially closed systems, namely the Lotka-Volterra model. This can be viewed as the simplest member of a hierarchical family of nested families of related models. We will progress through a series of models, that make increasingly elaborate and general assumptions about how each species' affects the other. These models are cast in a manner comparable to the

basic equation of demography, equation (1), separating births and deaths. In order of increasing complexity, these models are as follows:

I. Classical Lotka-Volterra predator-prey model (no direct density dependence)

$$\frac{dP}{dt} = P[abN - m]$$

$$\frac{dN}{dt} = N[r - aP]$$
(11)

We start with the simplest, classical Lotka-Volterra model, which for the predator separates births (or production) from deaths (or biomass loss). Here *a* is an attack rate, *aN* is the functional response of each predator to its prey, and *baN* is its numerical response (i.e., birth rate). The quantity *m* is the predator's death rate. The prey has an intrinsic birth rate, and death rate, with the difference between them being its intrinsic growth rate *r*, and it suffers an additional source of mortality from predation, of a magnitude *aP*. It is well-known that this model leads to cycles, but these cycles are neutrally stable, and so are delicately sensitive to small changes in model assumptions. However, the qualitative message of this simple model that there is a propensity to oscillate in natural enemy-victim systems applies robustly across a wide range of models. Moreover, for the purposes of our overview of the theory of natural enemy-victim interactions, the Lotka-Volterra model provides an entrée into a very interesting theoretical and empirical issue, having to do with the nature of the interaction between predators and prey.

II. Predator-prey models with nonlinear functional responses

$$\frac{dP}{dt} = P[bf(N, P) - m]$$
(12)
$$\frac{dN}{dt} = Nr - f(N, P)P$$

The quantity f(N,P) is the functional response of the predator to its prey, which in principle can depend upon both prey and predator density. Now, there can be emergent direct dependence in each species (the diagonal terms in *J*), and so the equilibrium can be locally stable, or unstable, depending upon the particular mathematical form for f(N,P).

A fundamental determinant of stability in predator-prey interactions is thus the quantity f(N,P), which is the functional response of C.S. Holling (1959), broadened to include possible dependencies of feeding rates on predator density. Model (12) assumes that the predator's numerical response (its birth rate) is directly proportional to the functional response. This fits

many examples and so is a reasonable assumption (Arditi and Ginzburg 1989, p. 316), though there are certainly counter-examples.

An ongoing debate about a key component of natural enemy-victim theory

Many robust debates in natural enemy-victim ecology arise from different views about how one relates detailed mechanisms of interactions at the level of individuals between predators and their prey to the functional and numerical responses expressed in (12). It has been almost fifty years since Holling highlighted the importance of the functional and numerical responses in predation, so it is striking that there is still lively disagreement about this rather basic issue, leading to the verbal equivalent of fisticuffs in the ecological literature.

Prey-dependence

In one corner of the ring, we have proponents of the 'classical' perspective, in which the predator is strictly food- or prey-limited; thus, the functional response depends solely on prey numbers, f(N), and the numerical response depends in turn purely on the functional response, g(f(N)). For a fixed prey density, the total rate of predation upon the prey should then increase linearly with predator density.

Holling (1959) noted that a very general feature of predator feeding responses is that they asymptote with increasing prey numbers (his Type 2 functional response), but sometimes accelerate over a range of low densities (the Type 3 functional response) In the former case, this implies that that the per capita effect of predation declines with increasing prey abundance, so $\partial f_N / \partial N < 0$. This is destabilizing, and without any other sources of density dependence, the equilibrium is locally unstable. Murdoch and Oaten (1975) show that for the Type 3 response to be stabilizing, (when the predator numerical response is proportional to the functional response) it is not enough for the functional response to be accelerating, but to do so with sufficient magnitude: the precise stability condition is

 $\partial f_N / \partial N | * > f_N^* / N^*.$ (13)

Much of the literature on functional responses has explored how behavioral and physiological mechanisms in the predator (e.g., adaptive foraging behaviors) can influence the form of the functional response (Whelan and Schmidt 2007), and hence the stability properties of predator-prey interactions. It is important to recognize that prey behavior can also influence the shape of

the functional response (Brown and Kotler 2007). For instance, if prey actively hide in refuges from predation, but refuges are in limited supply, there can be direct density dependence via the functional response emerging from competition for refuges (Holt 1987). At low prey density, most prey find refuge, and so per capita mortality from predation is low. As prey increase, more are forced out of the refuges, so mortality rates rise. This translates into an accelerating functional response for the predator, and a trace of J that is negative. Refuges are thus broadly stabilizing (Murdoch and Oaten 1975).

Predator dependence

In the other corner of the ring, we have ecologists who challenge the hegemony of preydependence, for instance by asserting that the simplest assumption is that there is a law of diminishing returns, such that the resources available per consumer declines with increasing consumer numbers (Ginzburg and Colyvan 2004). The assumption that functional responses depend solely on prey numbers sometimes does hold. But often this is not true, for instance because predators interfere with each other while searching for prey (Hassell and Varley 1969, DeAngelis et al. 1975, Free et al. 1977). Skalski and Gilliam (2001) surveyed a number of examples, and found that in the majority of cases studied, strong effects of predator density upon the functional responses were observed (though caution has to be made to compare with care alternative hypotheses to direct predator interference, such as prey behavioral responses and spatial processes, Sarnelle (2003)).

Interference among predators has both stabilizing, and destabilizing, effects. In terms of the formalism presented above, one element in the trace of *J* becomes negative because of such density dependence; if this effect is sufficiently strong it may contribute to stability. However, if predator dependence greatly reduces per capita consumption at high predator numbers, the total impact of the predator on the prey is constrained, which may make it harder for the predator to keep up with a growing prey population. Basically, strong interference makes natural enemy-victim systems more dominated by "bottom-up" forces (Arditi et al. 2004).

Conversely, it should be noted that in some circumstances, increased predator abundance may *facilitate* attacks. This is expected when predators forage in groups, but it can also occur in less obvious situations. Miller et al. (2006) for instance reports that the number of Canada goose nests attacked per bald eagle, increased with eagle density. They suggest this might occur

because goose defense strategies become much less effective, at higher predator numbers. A prey individual that escapes by scuttling away from one predator, at high predator densities may simply move into the path of another predator. Alternatively, grouping behavior by prey may reduce predation and stabilize predator-prey interactions (Fryxell et al. 2008).

A variety of functional forms have been proposed that display predator-dependence. One that has sparked much controversy is "ratio dependence" (Arditi and Ginzburg 1989), where the functional response depends upon the amount of prey per predator, *N/P*, i.e., *f*(*N/P*), rather than just *N*, as in *f*(*N*). In the Lotka-Volterra model, the functional response is a *N*, and the total mortality imposed by predators upon their prey is *aNP*. With ratio-dependent predation, the functional response is instead *a*(*N/P*), and the total mortality inflicted upon prey by predators is *a*(*N/P*)*P* = *aN*, and so the per capita rate of predation experienced by the prey is a constant, *a*, independent of predator abundance. The truth is likely bracketed by these, as represented for instance with the Beddington-DeAngelis functional response (Beddington 1975, DeAngelis et al. 1975):

$$\frac{aN}{1+ahN+iP} \quad (14)$$

where *a* and *N* are as above, *h* is a handling time, and *i* a measure of predator interference. Using (14), the total predation experienced by the prey at low predator and prey numbers is f(N,P)P = a'RP, as in the Lotka-Volterra model, but at high predator numbers (for fixed prey abundance) the total amount of predation becomes a'R/i, and so becomes independent of predator abundance.

There are empirical examples that fit both prey-dependent and predator-dependent scenarios, or that seem to lie in between (e.g., Schenck et al. 2005). In systems that are spatially well-mixed and with continuous reproduction, prey-dependent models often perform reasonably well (e.g., lake plankton, Carpenter et al. 1993; bacteria and virus in a chemostat, Bohannan and Lenski 1997). But in systems with spatial heterogeneity, pulsed reproduction, and other complexities, alternative formulations may be required, and ratio dependence may at times provide a reasonable approximation. Vucevitch et al. (2002) argue that on Isle Royale, the interaction between wolves and moose reflects features of both prey and ratio-dependent models, and Jost et al. (2005) in re-examining the data suggest the patterns even tilt toward the latter. Among theoreticians, the majority opinion nonetheless appears to be that prey-dependence

should be the starting position in developing theory, with specific mechanisms that can lead to predator dependence incorporated, as needed. This indeed tends to be my own stance, but it should be emphasized that some ecologists at present strongly disagree. I refer the reader to the thoughtful review of Abrams and Ginzburg (2000) that lays out some of the elements of agreement and disagreement on both sides of this issue. There is thus an important lack of consensus, on a fundamental issue at the very heart of natural enemy-victim theory.

What is the natural generalization of the single-species exponential growth model to interacting natural enemy and victim species? I think most ecologists would say -- the Lotka-Volterra model (the left side of (11)). The trophic term in this model in effect mimics the law of mass action for bimolecular reactions in liquid media chemistry, where the rate of the reaction is proportion to the multiple of the concentrations of the two interacting "species" (atoms, molecules, ions). This would seem at first glance to be the simplest way one could incorporate trophic interactions into a term that could be spliced onto equations for exponential growth for each species. To play devil's advocate, I would like to consider the possibility instead that maybe the appropriate generalization is a ratio-dependent model (I am here motivated by a discussion in Ginzburg and Colyvan 2004). Recall from above that the essential hallmark of exponential growth is that a population either goes extinct (if r<0), or grows to infinity (if r>0). So maybe the comparable model to exponential growth, for a trophic interaction, should also have these dynamical outcomes. The Lotka-Volterra model has (neutrally stable) *bounded* growth, so does not display this behavior. But consider the following model with "pure" linear ratio-dependent predation:

$$\frac{dN}{dt} = rN - a\left(\frac{N}{P}\right)P = (r - a)N$$

$$\frac{dP}{dt} = -mP + ba\left(\frac{N}{P}\right)P = -mP + baN$$
(15)

The quantity *a* is the *total* amount of predation imposed upon the prey population, which (because of ratio dependence) is independent of the number of predators present. If r < a, then both *N* and *P* decline deterministically to extinction. If r > a, both *N* and *P* grow to infinity. Deterministic extinction of both predator and prey, or unbounded growth, are found in a wide range of more general ratio-dependent models where the functional response is described by some f(N/P) (Jost et al. 1999; Lev Ginzburg, pers. comm..)

So a ratio-dependent model such as the Arditi-Ginzburg model (15) has the same dynamical outcomes as does exponential growth, and so one might argue that from an abstract point of view it actually may be the natural generalization to trophic interactions of exponential growth for a single population, rather than the Lotka-Volterra model. In the real world, of course, populations do not continue to grow to infinity, so the exponential model, and the ratio-dependent model clearly must break down at high abundances. But species really do go extinct, and a model of exponential decline is a reasonable starting point for understanding that process. And predators really do sometimes over-exploit their prey and then go extinct themselves, so the prediction of extinction by itself is not a fatal weakness of ratio dependence (Ackakaya et al. 1995, Ginzburg and Colyvan 2004). Lev Ginzburg would argue that this general line of reasoning provides a rationale for why ratio-dependence, rather than Lotka-Volterra-like preydependence, should provide the logical basis for developing predator-prey theory (L. Ginzburg, pers. comm.).

I should stress that these reflections are philosophical in nature, and are not necessarily a compelling argument in favor of using ratio-dependent models as the routine starting basis for trophic theory (e.g., for food webs; Abrams 1994). A number of authors (e.g., Berryman 1992, Getz 1984, 1999, McKane and Drossel 2006) do argue that the logical starting point of predatorprey theory is not the Lotka-Volterra model (as suggested by the hierarchical set of models of (11) through (12)), but alternative formulations such as the ratio-dependent model (15) The jury is still out on this very important issue (e.g., see the recent exchange between Fussman et al. 2007 and Jensen et al. 2007), and ratio-dependent models are certainly consistent with a number of widely reported patterns, such as equilibrial abundances in food changes along productivity gradients. As noted above, models that take the form of (2) assume spatial homogeneity, no seasonality, etc., and so circumvent some of the rationales suggested in the literature for ratio-dependence. In any case, all these simple models are in my view probably what Quine (1960) would call "limit myths" about nature, rather than representations of natural law.

The problem of scaling from individuals to populations

There are a number of important and subtle theoretical issues related to the debate over ratio dependence that are not yet resolved. The functional response is an attribute of individual predators. Models such as (2) and (11) are at the population level. So much of the argument

revolves around the most appropriate way to scale up from individual processes to population and community dynamics, which operate at different spatial and temporal scales (Arditi and Ginzburg 1989, Abrams and Ginzburg 2000). The entire issue of how one constructs an ecological theory that coherently embodies scaling among different levels of ecological organization is a central challenge, one that goes well beyond the particular issues of natural enemy-victim theory (O'Neill et al. 1986). Moment closure techniques (Keeling et al., 2000; Englund and Leonn 2008) are one promising approach for connecting among scales in natural enemy-victim theory, and are revealing fresh and unexpected insights. For instance, the functional response which best describes predation at a regional scale may differ between increasing and decreasing phases of predator-prey cycles, because of shifts in the covariance of local predator and prey densities (Englund and Leonn 2008).

To me, one of the most important and interesting issues raised by this debate is that when there is predator-dependence, the total rate of predation (for a fixed prey density) does not increase linearly with predator density. In some systems, the relationship between total predation and predator abundance can differ among prey species, for the same predator species (Essington and Hansson 2005). With pure ratio-dependence, the total predation rate is in fact completely independent of predator abundance. Abrams (1993) argues that for many reasons (e.g., adaptive prey responses, see Sih chapter), one should not expect to see a linear relationship between predation and predator abundance, and might even at times observe a negative relationship. The relationship between total predation rate and predator abundance needs much more focused empirical study across a wide range of taxa and systems.

Alternative modes of density-dependence in predator-prey interactios.

Models (11) and (12) assumed no direct density dependence, within or between species, other than those mediated through the interspecific interaction. Adding such density dependence leads to the general formulation

III.
$$\frac{dP}{dt} = P[g(f(N,P),P) - m(N,P)]$$
$$\frac{dN}{dt} = Nr(N,P) - f(N,P)P$$
(16)

Now, we have direct density dependence in recruitment of the prey (e.g., due to competition for resources) via the term r(N,P), and in predator mortality in the term m(N,P) (e.g., due to

aggressive interference) due to causes extrinsic to the trophic interaction. In the consumer, direct density dependence can depress the numerical response below that expected from the functional response alone (e.g., as in the house mice feeding on beech seeds studied by Ruscoe et al. 2005). In general, adding extrinsic negative density dependence to continuous time predator models such as (2) stabilizes the interaction (Rosenzweig 1971, Deng et al. 2007).

In addition to such extrinsic density dependence, and density dependence via the functional response, several other modalities of density dependence may be at play in predatorprey interactions, contained in the expressions r(N,P) and m(N,P). For instance, prey recruitment may depend directly upon predator density, for reasons other than direct consumption, if prey reduce their foraging effort in response to the presence of predators (Brown and Kotler 2007). The functional response of the predator may then indirectly reflect resource levels for the prey, leading to potentially complex relationships between prey density and predator feeding rates. Or, in escaping predation, the prey may become exposed to alternative sources of mortality. Some of the most important directions in contemporary natural enemyvictim theory have to do with elucidating the consequences of changes in prey behavior and other traits in the face of predation (Bolker et al.2003, Sih chapter). Brown and Kotler (2007) suggest that 'the ecology of fear' can resolve classical conundrums in community ecology. For instance, when predators become common, herbivores often reduce their feeding rates (e.g., to become more vigilant), and in effect trade off reduced fecundity for increased survival (Creel and Christianson 2008). This both buffers herbivores, and relaxes consumption on plants. If prey vigilance increases with predator abundance, the feeding rate per predator should decline with increasing predator abundance. Approximate ratio-dependence thus may emerge from an intrinsically prey-dependent system, because of adaptive flexibility in the prey. Brown and Kotler (2007) go on to observe that "paradoxically, efficient predators produce intrinsic instability to predator-prey dynamics, while inefficient predators produce vulnerability to extrinsic variability in prey population numbers. Fear responses by prey can break this paradox. At low predator numbers, the predator can efficiently catch unwary prey. At higher predator numbers, the predator becomes less efficient as the prey become increasingly uncatchable." Thus, prey behavioral responses may be broadly stabilizing, in the sense of bounding trajectories away from boundaries.

In the predator component of the general model given by (16), two other extensions of the basic Lotka-Volterra model are included. The extra P found as an argument inside the function g expresses the fact that for a given rate of prey capture, predator birth rates may vary directly with their own density (e.g., because of competition for limited nest sites). Predator death rates (the m(N,P)) may also depend jointly on prey and predator abundance (e.g., higher feeding rates may lead to lower predator deaths, not just higher predator births). For instance, in warm-blooded vertebrates, if the feeding rate goes below a certain level, death may rapidly occur. A number of forms relating the predator numerical responses to the functional response have been suggested in the literature (Ramos-Jileberto 2005). It would be useful for ecologists in developing theories to disentangle distinct birth and death responses in addressing these relationships. More broadly, much more attention needs to be paid to characterizing the functional form of predator numerical responses to their prey.

THE SUBDOMAIN OF HOST-PATHOGEN DYNAMICS

As noted earlier, different subdomains of natural enemy-victim theory have had somewhat different theoretical traditions. Consider models for interactions between hosts and pathogens such as bacteria and viruses. A pathogen is a natural enemy (usually a parasitic microorganism) that lives and reproduces as a population inside the victim, its host. The mathematical theory of epidemiology provides a theoretical framework for exploring the persistence and epidemic behavior of pathogens, and of the impact of pathogens upon their hosts. In a tradition going back to Kermack and McKendrick (1927), and summarized in a magisterial fashion by Anderson and May (1991) and more recently by Keeling and Rohani (2008), one subdivides host populations into classes such as susceptible, infected, and recovered individuals, with densities *S*,*I*, and *R* respectively. For microparasites such as viruses and bacteria, this theoretical maneuver means that one initially ignores the dynamics of the pathogen within individual hosts, and so pathogen abundance is not tracked directly, but instead the infection status of the host.

For simplicity, and to be consistent with the natural enemy-victim model given by (2), I will assume that we can use a two class model of just *S* and *I*, so recovery by an infected host leads to readmission into the infected class. An *SIS* model has the following general form:
$$\frac{dS}{dt} = R - i(I,S) - mS + \gamma I$$

$$\frac{dI}{dt} = i(I,S) - m'I - \gamma I$$
(18)

Here, *R* is the rate of recruitment of fresh susceptibles into the host population (which may depend upon host population size), i(I,S) is the total rate of infection, *m* and *m*' are per capita mortality rates, and γ is a per capita rate of recovery. It is useful to briefly compare the predator-prey model (12) with (18). The "prey" here are susceptible hosts, and the "predator" are infected hosts. The total rate of infection is analogous to the total mortality imposed upon the prey by the predator, i.e., the functional response times predator density. It is also the numerical response of the "predator" to susceptible hosts. These analogies suggest the following analogues of Propositions 6-9 (the primes are to indicate the corresponding propositions above):

Proposition 7': The total rate of infection (new infected individuals, per unit time) increases with the density of susceptible individuals.

Proposition 8' The total rate of infection increases with the density of infected hosts.

Proposition 9': For a specialist pathogen, the infection dies out in the absence of susceptible hosts.

There is a new proposition that does not have such a ready analogoue in predator-prey interactions:

Proposition 10: Susceptible hosts may be generated from infected hosts (by recovery, or possibly birth)

The last proposition has an important implication in terms of population regulation and stability. In human epidemiological models, it is often assumed that infection has no impact upon host population size, either through death or fecundity. If recruitment R = bN, where N=I+S, and the mortality rates are m=m'=b, the host population stays at its initial size (Swinton et al. 2002, Box 5.1). This assumption requires that the bioenergetic and fitness impact of infection be trivial, relative to other factors limiting and regulating host numbers. Thus, the infectious disease by assumption is irrelevant to host population regulation.

By contrast, the pathogen may not only influence host mortality or fecundity, but be the sole factor regulating the host. If we replace *R* with a constant per capita birth term, *bS*, and assume that infected individuals give birth at rate *b*', and recover at rate γ , (18) becomes:

$$\frac{dS}{dt} = bS + (b' + \gamma)I - i(I,S) - mS$$

$$\frac{dI}{dt} = i(I,S) - m'I - \gamma I$$
(19)

Adding the above equations leads to

$$\frac{dN}{dt} = (b-m)S + (b'-m)I \tag{20}$$

Thus, a *necessary* condition for regulation of a host by a pathogen, when the host is otherwise unregulated, is that the birth rate (of healthy hosts) of infected individuals must be less than their own death rates (Holt and Pickering 1987).

Whether or not regulation actually occurs depends upon the form of the transmission function i(I,S), which corresponds formally to both the total mortality term in the prey equations and the numerical response in the predator equations of (11). Different functional forms for transmission correspond to different assumptions about the biology of transmission (McCallum et al. 2001). A very active area of infectious disease ecology is focused on refining models for transmission (e.g., Keeling 2005), for instance using network models (Brooks et al. 2008). Hochberg (1991) provides a treatment of model (19), for general nonlinear models of transmission. There are two idealized forms of transmission -- density-dependent, and frequency-dependent -- as follows:

$$i(I,S) = \beta IS$$

$$i(I,S) = \frac{\beta IS}{N}$$
(21)

The former may be reasonable for airborne pathogens, but the latter is often more sensible for diseases where infected individuals come into contact with an approximately fixed number of potential hosts (e.g., sexually-transmitted diseases).

Note that with density-dependent transmission, equation (19) closely resembles the Lotka-Volterra predator-prey model, but with extra terms (b'I and γI) representing recruitment into the susceptible portion of the population, from the infected portion of the population. So even if susceptible hosts go to zero, the population can recover, because some are regenerated from infected hosts. The disease can in this case regulate the host, although when this extra term is small, the system exhibits weakly damped oscillations (Figure 7.1); as this term approaches

zero, the model becomes identical to the classical, Lotka-Volterra predator-prey model, and so has neutrally stable cycles.

Density-dependent transmission implies a threshold host density for disease invasion and persistence, and also permits host regulation by parasitism. By contrast, if the frequency-dependent term (the second expression in (21) is introduced into (19), there is no longer a threshold host population size, and the disease on its own is not able to regulate the host to a stable equilibrium (Getz and Pickering 1983). A mixture of density-dependent and frequency-dependent transmission can imply infection dynamics where a pathogen invades, increases in prevalence, and in so doing drives both the host and itself extinct (Ryder et al. 2007).

This difference between the outcomes of parasitism for host regulation, depending on the functional form of transmission, corresponds, if not in detail, at least in some features, to the difference between prey-dependent and ratio-dependent functional responses in general predatorprey theory. Frequency-dependent transmission can be written as follows (recalling that N=I+S):

$$\frac{\beta IS}{N} = \left(\frac{\beta S / I}{1 + S / I}\right) I . \quad (22)$$

The term in parentheses in (22) is the 'functional response' for the pathogen, i.e., the rate of new infections, per infected host. This is expressed as a function of the ratio of healthy hosts ("prey") to infected hosts ("prey"), and so frequency-dependent disease transmission can be viewed as a form of ratio-dependence. A functional form for trophic linkages that is still controversial in the subdomain of predator-prey theory is a run-of-the-mill assumption in another subdomain of natural enemy-victim theory, namely host-pathogen theory.

Note that if we look at model (19) with frequency-dependent transmission, in the limit when the infection is rare, the equations are approximately as follows:

$$\frac{dS}{dt} = bS - \frac{\beta IS}{N} - mS \approx (b - m)S$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - m'I \approx (\beta - m')I$$
(23)

This is simply a pair of equations for two populations growing exponentially. This matches the "devil's advocate" position about ratio-dependence and exponential growth, noted above.

Reflections on the differences between host-pathogen and predator-prey subdomains

At first glance, one sharp difference between the infectious disease model (19) and predator-prey models (as in (11)) is that recruitment into the susceptible class of the population can come from the infected class. But recall that an abstraction in *SI*-style models is that one simply counts the pathogen as present or absent, rather than enumerating them within each host. Hosts, in effect, are patches (with their own dynamics) that are colonized by pathogens. A very active area of research focuses on accounting for within-individual host population dynamics of pathogens. The appropriate analogue may be found only by shifting scales. The predator-prey model that in some ways may be more truly analogous to an *SI* model is a metapopulation model, where predators and prey are in patches coupled by dispersal, and the state of each patch may be empty, prey alone, or predator together with prey (Nee et al. 1997). If prey in patches with predators can still produce propagules that can reach and colonize empty patches, the prey equation in the metapopulation model will have a term comparable to γI in the *SI* model (Holt 1997a).

Another difference between traditions in predator-prey and host-pathogen ecology is that the former has a robust sub-tradition of models grounded in bioenergetics (see above), whereas scant attention has been paid to resource relationships, energetics, and stoichiometry in hostpathogen dynamics. But there is increasing evidence that resources modulate many aspects of the interactions between hosts and pathogens (Smith and Holt 1996, Smith 2007, Hall et al. 2009), and I suspect that this seeming difference in emphasis will change in the near future.

I will mention one deep-rooted difference between host-pathogen systems, and predatorprey systems, that has to do with basic biology, rather than mathematical abstractions, and which has not yet been addressed in depth in the literature. Because pathogens and parasites are symbionts, living intimately with their hosts, they can evolve into mutualists (Hochberg et al. 2000). Some taxa can even be facultative, parasitic in one setting, but a mutualist in others. This part of the domain of natural enemy-victim interactions shades into the domain of interspecific mutualisms.

TOWARDS MULTISPECIES NATURAL ENEMY - VICTIM THEORY

We have focused on the interaction between a specialist natural enemy and its victim. Other species may actually be present and have an impact on the two focal species, but it is assumed all such community interactions outside these two species can be swept into model parameters or functional forms (e.g., density-independent mortality rates for a predator might reflect its use of

alternative food resources than the focal prey species); what is important in this assumption is that the abundances of these other species do not need to be explicitly tracked, and do not provide alternative modes of dynamical feedbacks. Understanding multi-species interactions is an important and vibrant area of natural enemy-victim theory (and indeed has been the theme of most of my own papers in this area), shading into other domains in ecology (e.g. food web and network theory, competition theory). The assumptions one makes about the basic interaction between a single natural enemy species and a single victim species (e.g., prey-dependent vs. predator-dependent functional responses) can have profound consequences for the dynamics of complex food webs (e.g., Drossell, et al. 2001; McKane and Drossel 2006). I will here merely make a few remarks about how the propositions and models presented above need to be modified when there are multiple species present.

Consider again Proposition 3. There are complexities and subtleties here. First, (even for a single victim species) the statement applies *on average*. In spatially and temporally varying environments, or if the consumer can exploit multiple types of victims, it may not (and probably does not) pertain to each and every act of consumption. A generalist predator sustained by many prey species could include in its diet a few prey types that are a net loss, per act of consumption (see Holt 1983). Likewise, Proposition 9 surely holds for a generalist, only if *all* of its victims are absent; the natural enemy may simply equilibrate at a different (but non-zero) density in the absence of any particular victim species.

If one assumes that propositions 6 through 8 hold for multispecies natural enemy-victim models, then this has implications for the domain of community theory. In Holt (1977) I showed that if one assumed that (a) a predator feeding on two prey types has a positive numerical response to an increases in the abundance of each (i.e., propositions 6 and 7), (b) the predator is food-limited, so has no direct density dependence; and (c) the system settles into an equilibrium, then an alternative prey should always depress the equilibrial abundance of a focal prey species. So (-,-) interactions in a prey trophic level can emerge via the impact of a higher trophic level, i.e. apparent competition, which can at times constrain coexistence. This is an example of how propositions and theories in one domain (basic natural enemy-victim theory) can provide building blocks for constructing theory in another domain (community theory; see also Appendix for some community-level implications of the debate about predator- vs. prey-dependent functional responses).

Consider the basic Lotka-Volterra predator-prey model (11). One can add a third species to this system (still only considering trophic interactions) in three ways: i) a second predator species, consuming a shared resource; ii) a second prey species, sharing a generalist predator, and iii) a species at another trophic level, i.e., a three-species food chain. In the first two cases, one nearly always observes the exclusion of one species by another, via exploitative competition (for i), or apparent competition (for ii). The Lotka-Volterra food web is dynamically unstable; the top level constrains the abundance of the middle level, and so the bottom level can grow, unchecked. By working through the hierarchy of models presented above (nonlinear functional responses, direct density dependence, etc.), one can systematically examine the implications of these generalizations for species coexistence and food chain (and ultimately food web) regulation and stability. For instance, direct density dependence in a predator can stabilize a predator-prey interaction. If there are two predators present competing for one shared prey species, one normally would expect competitive exclusion, but this may not occur because the superior competitor exerts strong density dependence on itself via interference, freeing resources for the inferior competitor. In effect, this conceptual protocol provides a structured approach for identifying mechanisms of coexistence and exclusion, key desiderata in the domain of community ecology, using as a platform propositions and models crafted in another domain, namely natural enemy-victim theory. In turn, some conundrums in one domain may be resolved by consideration of themes in another. Specialist predator-prey interactions tend to be unstable, but with multiple species present additional forces can operate (e.g., switching by generalist predators) which help stabilize any given pairwise interaction, or at least keep fluctuations within reasonable bounds. Moreover, the instability generated by natural enemy-victim interactions creates another dimension of temporal variability in the environment, which can in turn influence the coexistence of competing species.

CONCLUDING THOUGHTS

In conclusion, I should stress that there is still a great deal of creative work ongoing in natural enemy-victim theory, and a future synthesis might look rather different from what I have presented here. Some areas of this domain (e.g., plant-herbivore interactions) warrant much more theoretical attention. For others -- continuous time predator-prey theory, host-parasitoid theory, and infectious disease theory -- recent work has highlighted the importance of important

theoretical extensions for understanding issues of stability and persistence, well beyond the factors emphasized in the classical two-species theory summarized above. These can include for instance age and stage structure in the interacting species. Using a single number (density) to characterize a population may be entirely inadequate for most natural systems; models with additional variables describing the internal state (e.g., energy reserves) of both consumers and their victims might be necessary. Crucially, space, localized interactions, environmental heterogeneity, and dispersal all have profound effects on all natural enemy-victim interactions (Briggs and Hoopes 2004). Keeling et al. (2000) in an important paper showed for the Lotka-Volterra and Nicholson-Bailey models embedded in space that the localized covariance between predators and prey leads to equations which help unify understanding of seemingly disparate causes for stability, including localized competition among predators. There are increasing efforts made to deal with the consequences of stochasticity at the level of individuals, which can scale up to entire population; Chesson (1982) provides a prescient overview of some of these issues. Many pairwise natural enemy-victim interactions cannot be understood without careful attention to multispecies interactions in a broadcommunity context (Holt1997b), and indeed ecosystem feedbacks (Loreau and Holt 2004). Future attempts toward a unified theory of natural enemy-victim interactions will need to deal with all of these issues in a systematic manner.

Bioenergetic and stoichiometric techniques potentially provide powerful tools for refining natural enemy-victim models, going well beyond what I can do justice to here. Bioenergetics integrated with models of exploitation have provided important insights into food web structure and dynamics (Williams et al. 2007). An influential attempt to ground predatorprey dynamics in organismal biology by Yodzis and Innes (1992) expressed consumption rates and biomass loss rates with allometric relationships for production, metabolism, and maximum consumption rates, thus using physical principles to constrain models of trophic interactions. This work is an intellectual predecessor to recent work in the metabolic theory of ecology (Brown et al. 2004, Savage et al. 2004), which potentially provides a general conceptual framework for linking traits of individual organisms with population and community dynamic models.

Across a wide range of taxa, there is a relationship between natural mortality rates and body size and temperature (McCoy and Gilooly 2008). Natural mortality includes predation, and coupled with information on predator abundance, this relationship sets upper bounds on ambient attack rates. This approach potentially permits one to parameterize predator-prey models in a way that does not require measuring each parameter afresh in each new setting (e.g., Vasseur and McCann 2005). A consideration of the physical dimension of life naturally leads to a focus on how feeding rates are reflected in the internal states of organisms, such as energy and nutrient reserves, which can influence patterns of decline in starving populations, and the quality of offspring (via maternal effects) (Ginzburg and Colyvan 2004). I suspect that future attempts to craft a general theory of natural enemy-victim interactions will be much more heavily and fundamentally influenced by these physical perspectives, than just the light patina I have here provided.

And finally, another very importance direction for future natural enemy-victim theory is to incorporate Darwinian evolution, which in many taxa can occur on time-scales commensurate with population dynamics. Coevolutionary cycles can feed back onto population dynamics, and vice versa, so that understanding the dynamics of these systems requires paying close attention to the generation, maintenance and reshuffling of genetic variation in each interacting player. Differences among systems in genetic variation and architecture may prove to be as consequential in determining their dynamical behaviors, as the ecological factors such as functional responses I have emphasized here in this attempt at a synthetic overview of natural enemy-victim theory.

Acknowledgements

I thank Sam Scheiner and Mike Willig for their invitation to participate in this endeavor, Sam and two anonymous reviewers for very helpful and detailed comments, Vitrell Sherif for assistance with manuscript preparation, and the University of Florida Foundation for support. Literature cited

Abrams, P.A. 1992. Predators that benefit prey and prey that harm predators: unusual effects of interacting foraging adaptations. The American Naturalist 140:573-600.

- Abrams, P.A. 1993. Why predation rate should not be proportional to predator density. Ecology 74:726-733.
- Abrams, P.A. 1994. The fallacies of "ratio-dependent" predation. Ecology 75:1842-1850.
- Abrams, P.A. 1997. Anomalous predictions of ratio-dependent models of predation Oikos 80:163-171.
- Abrams, P.A. 2007. Habitat choice in predator-prey systems: Spatial instability due to interacting adaptive movements. The American Naturalist 169:581-594.
- Abrams, P.A. and L.R. Ginzburg. 2000. The nature of predation: prey dependent, ratio dependent, or neither? Trends in Ecology and Evolution 15:337-341.
- Akcakaya, H.R., R. Arditi, and L.R. Ginzburg. 1995. Ratio-dependent predation: an abstraction that works. Ecology 76:995-1004.
- Anderson, R.M. and M.P. Hassell. 1989. Predator-prey and host-pathogen interactions. pp. 147-196 in Ecological Concepts, J.M. Cherrett, ed. Blackwell.
- Anderson, R.M. and R.M. May. 1991. Infectious Diseases of Humans. Oxford University Press, Oxford, UK. 757 p.
- Arditi, R., J.-M. Callois, Y. Tyutyunov, and C. Jost. 2004. Does mutual interference always stabilize predator-prey dynamics? A comparison of models. Comptes Rendus: Biologies 327:1037-1057.
- Arditi, R. and L.R. Ginzburg. 1989. Coupling in predator-prey dynamics: ratio-dependence. J. Theor. Biol. 139:311-326.
- Beddington, J.R. 1975. Mutual interference between parasites or predators and its effect on searching efficiency. Journal of Animal Ecology 44:331-340.
- Begon, M., C.R. Townsend and J.L. Harper. 2006. Ecology: From Individuals to Ecosystems. Blackwell, UK.
- Berryman, A.A. 1999. Principles of Population Dynamics and Their Applications. Stanley Thornes.
- Bohannan, B.J.M. and R.E. Lenski. 1997. Effect of resource enrichment on a chemostat community of bacteria and bacteriophage. 78:2303-2315.

- Bolker, B., M. Holoak, V. Krivan, L. Rowe and O. Schmitz. 2003. Connecting theoretical and empirical studies of trait-mediated interactions. Ecology 84:1101–1114.
- Borer, E.T., C.J. Briggs and R.D. Holt. 2007. Predators, parasitoids, and pathogens: a crosscutting examination of intraguild predation theory. Ecology.
- Briggs, C.J. and M.F. Hoopes. 2004. Stabilizing effects in spatial parasitoid-host and predatorprey models: a review. Theoretical Population Biology 65:299-315.
- Brooks, C.P., J. Antonovics and T.H. Keitt. 2008. Spatial and temporal heterogeneity explain disease dynamics in a spatially explicit network model. The American Naturalist 172: 149-159.
- Brown, J.H., J.F. Gillooly, A.P. Allen, V.M. Savage, and G.B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771-1789.
- Brown, J.S. and B.P. Kotler. 2007. Foraging and the ecology of fear. Pp. 437-482 in Foraging: Behavior and Ecology. eds. D.W. Stephens, J.S. Brown, and R.C. Ydenberg. University of Chicago Press.
- Carbone, C. and J.L. Gittleman. 2002. A common rule for the scaling of carnivore density. Science 295:2273-2276.
- Caro, T. 2005. Antipredator Defenses in Birds and Mammals. University of Chicago Press, Chicago, IL.
- Carpenter, S.R., R.C. Lathrop, and A. Munoz-del-Rio. 1993. Comparison of dynamic models for edible phytoplankton. Can.J. Fish. Aquat. Sci. 50:1757-1767.
- Caswell, H. 2001. Matrix Population Models. 2nd ed. Sinauer, Sunderland, Ma.
- Caughley, G. and J.H. Lawton. 1981. Plant-herbivore systems. Theoretical Ecology, 2nd ed., R.M. May, ed. Sinauer Associates, Sunderland, MA.
- Crawley, M.J. 1983. Herbivory: The Dynamics of Animal-plant Interactions. Blackwell. 437 p.
- Crawley, M.J., ed. 1992. Natural Enemies: The Population Biology of Predators, Parasites and Diseases. Blackwell, Oxford.
- Creel, S. and D. Christianson. 2008. Relationships between direct predation and risk effects. Trends in Ecology and Evolution 23:194-201.
- Dawkins, R. 1982. The Extended Phenotype. W.H. Freeman, Oxford.
- DeAngelis, D.L., R.A. Goldstein, and R.V. O'Neill. 1975. A model for trophic interaction. Ecology 56: 881–892.

- Deng, B., S. Jessie, G. Ledder, A. Rand and S. Srodulski. 2007. Biological control does not imply paradox. Mathematical Biosciences 208:26-32.
- de Roos, A.M., L. Persson, and H.R. Thieme. 2003. Emergent Allee effects in top predators feeding on structured prey populations. Proc. Royal Soc. Lond. Series B-Biological Sciences 270:611-618.
- Diamond, J. 1992. Living through the Donner party. Discovery 13(3):100-107.
- Drossel, B. P.G. Higgs and McKane 2001. The influence of predator-prey population dynamics on the long-term evolution of food web structure. J. Theor. Biol. 208:91-107.
- Eldgredge, N. 1985. Unfinished Synthesis: Biological Hierarchies and Modern Evolutionary Thought. Oxford University Press.
- Essington, T.E. and S. Hanson. 2004. Predator-dependent functional responses and interaction strengths in a natural food web. Can. J. Fish. Aquat. Sci. 61:2215-2226.
- Englund, G. and K. Leonardsson. 2008. Scaling up the functional response for spatially heterogeneous systems. Ecology Letters 11:440-449.
- Ernest, S.K.M., et al. 2003. Thermodynamic and metabolic effects on the scaling of production and population energy use. Ecology Letters 6:990-995.
- Free, C.A., J.R. Beddington, and J.H. Lawton. 1977. On the inadequacy of
- simple models of mutual interference for parasitism and predation,
- J. Anim. Ecol. 36: 375–389.
- Fryxell, J.M., A. Mosser, A.R.E. Sinclair and C. Packer. 2007. Group formation stabilizes predator-prey dynamics. Nature 449:1041-1043.
- Fussmann, G. F., G. Weithoff, and T. Yoshida. 2005. A direct, experimental test of resource vs. consumer dependence. Ecology 86:2924–2930.
- Fussmann, G.R., G. Weithoff and T. Yoshida. 2007. A direct, experimental test of resource vs. consumer dependence: reply. Ecology 88:1603-1604.
- Getz, W. M. 1984. Population dynamics: a resource per-capita approach. J. Theor. Biol. 108: 623-644.
- Getz, W. M. 1993. Metaphysiological and evolutionary dynamics of populations exploiting constant and interactive resources: r-K selection revisited. Evol. Ecol. 7: 287-305.

- Getz, W. M. 1999. Population and evolutionary dynamics of consumer-resource systems. Pp. 194-231 In: McGlade, J. (ed.) Advanced Ecological Theory. Principles and Applications.Blackwell Science.
- Getz, W.M. and J. Pickering. 1983. Epidemic models: thresholds and population regulation. The American Naturalist 121:892-898.
- Ginzburg, L. R. 1998. Assuming reproduction to be a function of consumption raises doubts about some population predator-prey models. J. Anim. Ecol. 67: 325-327.
- Ginzburg, L. and M. Colyvan. 2004. Ecological Orbits: How Planets Move and Populations Grow. Oxford University Press, Oxford, U.K.
- Grassly, N.C. and C. Fraser. 2008. Mathematical models of infectious disease transmission. Nature Reviews Microbiology (published online 13 May 2008; doi:10.1038/nrmicro1845).
- Gutierrez, A. P. 1992. Physiological basis of ratio-dependent predator-prey theory: the metabolic pool model as a paradigm. Ecology 73: 1552-1563.
- Hall, S.R. et al. 2008. Is infectious disease just another type of predator-prey interaction? Pp. 223-241 In Ostfeld, R.S., F. Keesing, and V.T. Eviner, eds. Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems. Princeton University Press, Princeton, NJ.
- Hall, S.R., C.J. Knight, C.R. Becker, M.A. Duffy, A. J. Tessier, and C.E. Caceres. 2009. Quality matters: resource quality for hosts and the timing of epidemics. Ecology Letters 12:1181-28.
- Hassell, M.P. and G.C. Varley. 1969. New inductive population model for insect parasites and its bearing on biological control. Nature 223:1133–1137.
- Hassell, M.P. 1978. The Dynamics of Arthropod Predator-Prey Systems. Princeton University Press, Princeton, NJ.
- Hassell, M.P. and S.W. Pacala. 1990. Heterogeneity and the dynamics of host-parasitoid interactions. Phil. Trans. R. Soc. London B 330:203-220.
- Hassell, M.P. 2000. The Spatial and Temporal Dynamics of Host-Parasitoid Interactions. Oxford University Press, UK.
- Hochberg, M.E. and J.H. Lawton. 1990. Spatial heterogeneities in parasitism and population dynamics. Oikos 59:9-14.
- Hochberg, M.E. 1991. Non-linear transmission rates and the dynamics of infectious disease. Journal of Theoretical Biology 153:301-321.

- Hochberg, M.E., R. Gomulkiewicz, R.D. Holt and J.N. Thompson. 2000. Weak sinks could cradle mutualistic symbioses – strong sources should harbor parasitic symbioses. Journal of Evolutionary Biology 13:213-222.
- Holling, C. S. 1959. The components of predation as revealed by a study of small mammal predation of the European pine sawfly. Canadian Entomologist 91:293–320.
- Holt, R.D. 1977. Predation, apparent competition, and the structure of prey communities. Theoretical Population Biology 12:197-229.
- Holt, R.D. 1983. Optimal foraging and the form of the predator isocline. The American Naturalist 122:521-541.
- Holt, R.D. and J. Pickering. 1985. Infectious disease and species coexistence; a model of Lotka-Volterra form. The American Naturalist 126:196-211.
- Holt, R.D. and B.P. Kotler 1987. Short-term apparent competition. The American Naturalist 142:623-645.
- Holt, R.D. 1997a. From metapopulation dynamics to community structure: some consequences of spatial heterogeneity. Pp. 149-164 in Metapopulation Biology, eds. I. Hanski and M.E. Gilpin. Academic Press, San Diego, CA.
- Holt, R.D. 1997b. Community modules. Pp.333-350 in Multitrophic Interactions in Terrestrial Systems,eds. A.C. Gange and V.K. Brown. Blackwell Science, Oxford.
- Holt, R. D and M.E. Hochberg. 1998. The coexistence of competing parasites. II.Hyperparasitism and food chain dynamics. Journal of Theoretical Biology 193: 485-495
- Holt, R. D. 2002. Food webs in space: On the interplay of dynamic instability and spatial processes. Ecological Research 17:261-273.
- Hudson, P.J., A. Rizzoli, B.T. Grenfell, H. Heesterbeek, and A.P. Dobson. 2002. The Ecology of Wildlife Diseases. Oxford University Press.
- Inchausti, P. and L.R. Ginzburg. 2009. Maternal effects mechanism of population cycling: a formidable competitor to the traditional predator-prey view. Phil. Trans. Roy. Soc. B-Biological Sciences 364:1117-1124.
- Jensen, C. X. J., J.M. Jeschke, and L. R. Ginzburg. 2007. A direct, experimental test of resource vs. consumer dependence: comment. Ecology 88:1600-1602.
- Jost, C., O. Arino, and R. Arditi. 1999. About deterministic extinction in ratio-dependent predator-prey models. Bulletin of Mathematical Biology

- Jost, C., G. Devulder, J.A. Vuctich, R.O. Peterson and R. Arditi. 2005. The wolves of Isle Royale display scale-invariant satiation and ratio-dependent predation on moose. Journal of Animal Ecology 74:809-816.
- Keeling, M.J. 2005. Extensions to mass-action mixing. pp. 107-142 in K. Cuddington and B.E. Beisner, eds. Ecological Paradigms Lost: Routes of Theory Change. Elsevier.
- Keeling, M.J. and P. Rohani. 2008. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton, NJ. 368p.
- Keeling, M.J., H.B. Wilson and S.W. Pacala. 2000. Reinterpreting space, time lags, and functional responses in ecological models. Science 290:1758-1760.
- Kermack, W.P. and A.G. McKendrick. 1927. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. London A 115:700-712.
- Kolomogorov. A.N. 1936. Sulla Teoria di Volterra della Lotta per L'Esisttenza. Giorn. Instituto Ital. Attuari 7:74-80.
- Kooijman, S. A. L. M. 1993. Dynamic Energy Budgets in Biological Systems. Cambridge University Press.
- Kotler, B.P. and R.D. Holt. 1989. Predation and competition: the interaction of two types of species interactions. Oikos 54:256-260.
- Lafferty, K.D. and A.M. Kuris. 2002. Trophic strategies, animal diversity, and body size. Trends in Ecology and Evolution 17:507-513.
- Lockwood, D.R. 2008. When logic fails ecology. Quarterly Review of Biology 83:57-64
- Loreau, M. and R.D. Holt. 2004. Spatial flows and the regulation of ecosystems. The American Naturalist 163:606-615.
- Lotka, A.J. 1926. Elements of Physical Biology. Williams and Wilkins, Baltimore
- Marquet, P.A. 2002. Of predators, prey, and power laws. Science 295:2229-2230.
- May, R.M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton, NJ.
- May, R.M. and C.H. Watts. 1992. The dynamics of predator-prey and resource-harvester systems. pp. 458-475 in Natural Enemies, ed. M. Crawley.
- McCallum, H. N. Barlow and J. Hone. 2001. How should pathogen transmission be modeled? Trends in Ecology and Evolution 16:295-300.

- McCauley, E., W.A. Nelson and R.M. Nisbet. 2008. Small-amplitude cycles emerge from stagestructured interactions in Daphnia-algal systems. Nature 455:1240-1243.
- McCoy, M.W. and J.F. Gillooly. 2008. Predicting natural mortality rates of plants and animals. Ecology Letters 11:710-716.
- McKane, A.J. and B. Drossel. 2006. Models of food-web evolution. pp. 223-243 in Ecological Networkds: Linking Structure to Dynamics in Food Webs, eds. M. Pascual and J.A. Dunne. Oxford University Press.
- Miller, D.A., J.B. Grand, T.E. Fondell and M. Anthony. 2006. Predator functional response and prey survival: direct and indirect interactions affecting a marked prey population. Journal of Animal Ecology 75:101-110.
- Murdoch, W.W., C.J. Briggs, and R.M. Nisbet. 2003. Consumer-Resource Dynamics. Princeton University Press.
- Murdoch, W.W. and A. Oaten. 1975. Predation and population stability. Advances in Ecological Reseasrch 9:1-131.
- Nee, S., R.M. May and M.P. Hassell. 1997. Two-species metapopulation models. Pp. 123-147 in Metapopulation Biology, I. Hanski and M. Gilpin, eds. Academic Press, London.
- Nicholson, A.J. and V.A. Bailey. 1935. The balance of animal populations. Proc. Zool. Soc. Lond. 1:551-598.
- Nunney, L. 1985. The effect of long time delays in predator-prey systems. Theoretical Population Biology 27:202-221.
- O'Neill, R.V., D.L. DeAngelis, J.B. Waide and T.F.H. Allen. 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton, NJ.
- Ostfeld, R.S. and F. Keesing. 2000. Biodiversity and disease risk: the case of Lyme disease. Conservation Biology 14:722-728.
- Ostfeld, R.S., F. Keesing, and V.T Eviner. 2008. Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems. Princeton University Press, Princeton, NJ.
- Owen-Smith, N. 2002. Adaptive Herbivore Ecology: From Resources to Populations in Variable Environments. Cambridge University Press. 374 p.
- Owen-Smith, N. 2005. Incorporating fundamental laws of biology and physics into population ecology: the metaphysiological approach. Oikos 111:611-615.

- Paine, R.T.. 1966. Food web complexity and species diversity. The American Naturalist 100:65-75.
- Paine, R.T. 1980. Food webs linkage, interaction strength and community infrastructure- the 3rd Tansley Lecture. Journal of Animal Ecology 49:667-685.
- Quine, W.V. O. 1960. Word and Object. New York: MIT Press and John Wiley and Sons.
- Ramos-Jiliberto, R. 2005. Resource-consumer models and the biomass conversion principle. Environmental Modelling and Software 20:85-91.
- Randolph, S.E., et al. 2002. The ecology of tick-borne infections in wildlife reservoirs. pp. 119-138 in Hudson, P.J., A. Rizzoli, B.T. Grenfell, H. Heesterbeek, and A.P. Dobson. The Ecology of Wildlife Diseases. Oxford University Press.
- Rosenzweig, M.L. 1971. Paradox of enrichment: Destabilization of exploitation ecosystems in ecological time. Science 171:385-387.
- Rosenzweig, M.L. and R.H. MacArthur. 1963. Graphical representation and stability of predatorprey interactions. The American Naturalist 97:209-223.

Royama, T. 1992. Analytical Population Dynamics. Springer.

- Ruscoe, W.A., J.S. Elkinton, D. Choquenot and R.B. Allen. 2005. Predation of beech side by mice: effects of numerical and functional responses. journal of Animal Ecology 74:1005-1019.
- Ryder, J.J., M.R. Miller, A. White, R.J. Knell, and M. Boot. 2007. Host-parasite population dynamics under combined frequency- and density-dependent transmission. Oikos 11:2017-2026.
- Sarnelle, O. 2003. Nonlinear effects of an aquatic consumer: causes and consequences. The American Naturalist 161:478-496.
- Savage, V.M., J.F. Gillooly, J.H. Brown, G.B. West and E.L. Charnov. 2004. Effects of body size and temperature on population growth. The American Naturalist 163:429-441.
- Scheiner, S.M. and M.R. Willig. 2008. A general theory of ecology. Theoretical Ecology 1:21-28.
- Schenk, D., L.-F. Bersier and S. Bacher. 2005 An experimental test of the nature of predation: neither prey- nor ratio-dependent. Journal of Animal Ecology 74:86-91.

- Schoener, T.W. 1986. Overview: Kinds of ecological communities ecology becomes pluralistic. Pp. 467-479 in Community Ecology, eds. J. Diamond and T.J. Case. Harper & Row, New York.
- Skalski, G.T. and J.F. Gilliam. 2001. Functional responses with predator interference: viable alternatives to the Holling Type II model. Ecology 82:3083-3092.
- Smith, V.H. 2007. Host resource supplies influence the dynamics and outcome of infectious disease. Integrative and Comparative Biology 47:310–316.
- Smith, V. H. and R.D. Holt. 1996. Resource competition and wihtin-host disease dynamics. Trends in Ecology and Evolution 11:386-389.
- Sterner. R.W. and J.J. Elser. 2002. Ecological Stoichiometry: The Biology of Elements form Moleculres to the Biosphere. Princeton University Press, Princeton, NJ.

Taylor, R.J. 1984. Predation. Chapman & Hall, London.

- Thrall, P.H., A. Biere and M.K. Uyenoyama. 1995. Frequency-dependent disease transmission and the dynamics of the Silene-Ustilago host-pathogen system. The American Naturalist 145:43-62.
- Turchin, P. 2001. Does population ecology have general laws? Oikos 94:17-26. Turchin, P. 2003.
- Vasseur, D.A. and K.S. McCann. 2005. A mechanistic approach for modeling temperaturedependent consumer-resource dynamics. The American Naturalist 166:184-198.
- Volterra, V. 1926. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Mem. del Acad. Lincei 2:31-113.
- Volterra, V. 1931. Lecons sur la Mathematique de la Lutte pour la Via. Marcel Brelot, Paris.
- Whelan, C.J. and K.A. Schmidt. 2007. Food acquisition, processing, and digestion.Pp. 141-172 in Foraging: Behavior and Ecology. eds. D.W. Stephens, J.S. Brown, and R.C. Ydenberg. University of Chicago Press.
- Williams, R.J., U. Brose and N.D. Martinez. 2007. Homage to Yodzis and Innes 1992: Scaling up feeding-based population dynamics to complex ecological networks. Pp. 35-52 in From Energetics to Ecosytems: The Dynamics and Structure of Ecological Systems. Springer, Dordrecth, Netherlands.
- Wilmshurst, J.F., J.M. Fryxell and P.E. Colucci. 1999. What constrains daily intake in Thomson's gazelles? Ecology 80:2338-2347.

- Wilmshurst, J.F., J.M. Fryxell and C.M. Bergman. 2000. The allometry of patch selection in ruminants. Roc. R. Soc. Lond. B 267:345-349.
- Yodzis, P. and S. Innes. 1992. Body size and consumer-resource dynamics. The American Naturalist 139:1151-1175.

Figure 7.1. An example of damped oscillations in a SI host-pathogen interaction, following a perturbation. The model is (19) in the main text, and transmission is assumed to be density-dependent (β SI). The parameters are b=1, b'=0, m=m'=0.5, γ =0.1, β =2. The infection is assumed to completely suppress host fecundity. A small amount of recovery by infected hosts provides a weakly stabilizing effect, which slowly damps the oscillations. (Solid line= S, healthy hosts; dashed line = I, infected hosts).



APPENDIX

My own interest in natural enemy-victim interactions stems from reflections on how this theory informs our understanding of community structure. The debate between prey-dependent and predator-dependent functional responses discussed in the main text has important implications for understanding the impact of predators on species coexistence. In Holt (1977), and subsequent publications, I have argued that alternative prey species can experience a negative (-,-) interaction, mediated through their shared predator (apparent competition). This can arise because of predator behavior (Holt and Kotler 1987), or because of the predator's numerical response – if sustaining more predators implies stronger predation pressure upon a focal prey species, which if if has a low recruitment rate can be excluded from a local community. Strong predator dependence in the functional response potentially weakens the indirect effect of one prey species upon another and thus reduces the potential role of apparent competition in prey community organization.

Consider a community where a generalist predator is at equilibrium with a resident prey (species 1), and the interaction between them is described by the Beddington-DeAngelis functional response. Because of predator interference, the resident prey will often equilibrate at higher abundance, than would otherwise be the case. Assume a second prey species (species 2), is introduced at low numbers, during community assembly, and that the rate of predator attack is also described by the same functional response. The rate of growth of this species when rare is given by

$$\frac{dN_2}{dt} = N_2(r_2 - a_2 \frac{P^*}{1 + a_1 h_1 N_1^* + iP^*}).$$

There are three distinct causal pathways by which predator interference influences the strength of apparent competition between the resident and invading prey species. First, interference directly reduces the per predator attack rate (this is the term iP^* in the denominator). Second, because the resident prey are more abundant (since they are experiencing fewer attacks, per predator, there is an additional reduction in attacks upon the invader, via the functional response (this is the term $a_1h_1N_1^*$ in the denominator). Finally, interference among predators can lower the abundance of predators sustained by the resident prey (the P^* term in the

numerator), and so weakens the apparent competition effect. A particularly extreme form of predator dependence is ratio-dependence. So issues in one domain, having to do with how one conceives of the interaction between a natural enemy and its victim, can have major consequences for how one conceives of phenomena and processes in another domain.

Chapter 8: The Metacommunity Concept and its Theoretical Underpinnings Mathew A. Leibold

A metacommunity is a set of local communities that are connected by the dispersal of at least one of the component species ((Hanski 1991; Leibold *et al.* 2004)). The theory of species interactions at local scales is fairly well developed (see Holt Chapter 7) and serves as the focus of much of what we know about in community ecology. Metacommunity ecology seeks to take this understanding and use it to consider ecological dynamics at larger spatial scales by considering how dispersal among local communities affects ecological dynamics. Thus, the concept is a useful way to integrate what we know about spatial dynamics in community ecology and because it expands the field of community ecology to larger spatial scales in a couple of important ways.

First there is the idea that dynamics that occur at one place (a single local community) are not necessarily independent from those that occur at other places. Although spatial effects have been thought about for a long time (e.g.,(Skellam 1951)) the metacommunity concept allows us to consider a wide array of ways in which this may occur.

Second, there is an implicit idea of multiple scales and the idea that somewhat different or additional ecological principles apply at larger scales than the local community. Thus the concept is one way to address how ecological processes may determine patterns of biodiversity and composition at scales that are larger than the local scale and address questions about 'mesoscale' ecology (Holt 1993b; Ricklefs 1993). The important idea here is that it is possible to derive predictions about the attributes of biotic assemblages at these meso-scales that are based only on ecological principles. Clearly there are yet larger scales where historical and biogeographic effects are important but the metacommunity concept may allow us to understand community structure at a wide range of spatial scales that fall short of this larger scale and to explore how ecological dynamics may interact with historical biogeographic ones at scales larger than that (e.g. (Cornell & Lawton 1992).

There is however, an important distinction to be made between the 'metacommunity concept' and 'metacommunity theory'. As a concept, 'the metacommunity' is proving to be stimulating and useful way to organize approaches to multiscale ecology (Figure 8.1). However,

metacommunity theory is still in its infancy and consists of several different sets of ideas and model formulations. This chapter seeks to describe and evaluate a 'constitutive theory' of metacommunities. The heterogeneity of approaches to metacommunities makes a philosophical discussion of such a theory difficult and bound to be incomplete. Furthermore, it seems likely that there are possible novel processes and approaches to metacommunities that will arise and not be adequately described by any current effort at describing a 'constitutive theory'. Such constitutive theory is thus likely to change substantially in the future as thinking about metacommunities develops further.

A well developed 'constitutive theory' in ecology should, to some degree, have the following features (Scheiner & Willig 2005):

- a domain that describes the phenomena and principles involved.

- a set of propositions that elucidate understanding of the domain.

- a set of more precise models that link this theory to data

I outline how I understand these aspects of metacommunity ecology theory below and then discuss how this helps elaborate a more sophisticated view of the metacommunity concept as a whole.

THE DOMAIN OF METACOMMUNITY THEORY

The domain of metacommunity ecology is to explain how ecological communities vary in space. This domain is very similar to that of community ecology in general but focuses particularly on how communities vary from one place to another rather than on the structure of individual communities. There are numerous ecological features that form the focus of metacommunity ecology that describe or follow from understanding how the distribution and abundance of interacting organisms is regulated by ecological processes and depend on dispersal. Some examples include:

- patterns of species composition (which species with which traits are found in communities) in space, in relation to environmental conditions, and dispersal
- patterns of species richness (how many species are found in communities) in relation to space, environment and dispersal

 consequent effects of composition, diversity and dispersal on other attributes of communities such as food web structure, stability, and ecosystem features (e.g. productivity).

It may be useful to focus discussion of the domain by pointing out that all of these effects largely boil down to understanding how dispersal and other ecological processes affect how community composition varies in relation to the various landscape features. A very crude way of studying this empirically is by variation decomposition of community composition into purely environmental, purely spatial, spatio-environmental colinearity, and residual effects and there has been a rapidly burgeoning literature on such patterns to try to interpret metacommunity dynamics in natural and experimental systems (e.g. (Cottenie 2005; Beisner *et al.* 2006). While more sophisticated ways of thinking about composition are clearly needed, one might argue that the central goal of metacommunity ecology is to explain them and explore the consequences that follow from this variation.

In addition however, metacommunity ecology also addresses similar questions at spatial scales that are larger than the local community. Thus it can address questions about the composition and diversity of sets of local communities. This greatly expands the utility of the theory because it can address novel patterns. Examples include patterns related to macroecology and to paleoecology where data often represent composite data aggregated over multiple spatial samples. In paleoecology this happens as a natural result of taphonomic processes and in macroecology this happens as a result of data extraction from range maps etc. At these larger spatial scales, metacommunity ecology also begins to share a domain with evolutionary biology and historical biogeography since these processes also explain some aspects of variation in composition and biodiversity.

THE PROPOSITIONS OF METACOMMUNITY THEORY

Metacommunity ecology combines classical community ecology (i.e. community ecology of closed communities) with dispersal. The central propositions of metacommunity ecology thus include all the propositions that apply to community ecology in combination with a set of additional propositions about the role of dispersal.

The most important and characterizing propositions of metacommunity theory focus on four different ways that dispersal among local communities can affect their dynamics:

- 1) Dispersal contributes to local community assembly by serving as a source of colonists for species that were previously absent. Ignoring how species originally came to exist in these communities to begin with, in the absence of dispersal, local communities can only change in species composition by extinctions or in-situ speciation. At such local scales, speciation would be considered to be sympatric and while this can occur it is thought to be less important than allopatric speciation. Thus community assembly is severely constrained if dispersal is absent. If we consider that local environmental conditions likely change (e.g. via succession, perturbations, etc.) it seems unlikely that species composition would be likely to track environmental conditions as well and if these cause numerous extinctions, local diversity is also likely low. The only ways that biotas could respond to such changes would be via in situ evolution and while this may be important, it would not produce associations between species composition and local environmental conditions. Thus, the observation that there are such associations indicates that this proposition is important.
- 2) Dispersal can generate population 'mass effects' (also referred to as 'source-sink' effects) between communities that differ in local fitness (Amarasekare & Nisbet 2001; Mouquet & Loreau 2002). These effects involve the maintenance of artificially high populations in some patches due to immigration from other patches where the same species is highly productive. An important consequence is that such dispersal can allow a species to exist in a local community where it cannot have a self-sustaining population because that population's existence is dependent on continuous immigration from other patches where that species has high fitness; these are the 'sink' populations. Alternatively, dispersal also involves emigration so that dispersal can also lead to extinctions of species in highly productive habitats that would otherwise be self-sustaining. These effects are often likely influenced by adaptive dispersal that minimizes these effects but this need not always to be so. One important consequence of this role for dispersal is that it may lead to reduced associations between local environmental conditions and species composition. This role for dispersal instead implies that local communities that are close together in space may share species independently of whether they are similar in local environments.
- Heterogeneity in dispersal among species can allow for novel ways of coexistence in a metacommunity. The most obvious case for this is the 'colonization-competition' trade-

off hypothesis. In dispersal limited systems (where colonization limits community assembly) with a regular frequency of local extinctions, a poor competitor can coexist in the metacommunity with superior competitors that always exclude it locally if it can colonize communities that have had recent extinctions ((Levins & Culver 1971; Hastings 1980)). Coexistence of species with different dispersal rates is also possible under more complex scenarios, including cases where source-sink relations are present (Snyder & Chesson 2004) but their dynamics is still poorly understood.

4) Metacommunities can have 'spatial feedbacks' due to a complex suite of additional effects that recurrent dispersal between patches can have (Holt 1984), pers. comm.). Although this proposition is still very poorly developed, the basic idea here is that dispersal may be important in providing temporally lagged feedback effects on local communities via residence in other local communities and back migration into these communities.

While these four propositions are distinctive to metacommunity ecology, other more conventional propositions are just as important. These include:

- 5) Within local communities organisms belonging to different species interact via a set of interactions involving both direct and indirect effects. Direct interactions include competition (reciprocally negative or -/- interactions), exploitation (-/+), mutualisms (+/+), commensalisms (0/+) and interactions involving one way negative effects (0/-). Indirect effects involve chains of such interactions that may also be complex by being non-linear or interactive (Levins 1975). *This corresponds to the second 'fundamental principle' of Scheiner and Willig (2005 (Scheiner & Willig 2005))*. Most of the theoretical work to date has focused only on competition (-/-) interactions but there is some emerging work on food webs (+/-). Future work should explore the full scope of interactions.
- 6) Interactions in local communities are context dependent in that they are affected by the local environmental abiotic conditions (the environmental template) and by the existence of other species in the community (reviewed in (Chase 2003)). *This corresponds to the fourth 'fundamental principle' of Scheiner and Willig (2005) (Scheiner & Willig 2005)*. A large number of models concentrate on cases where local communities exist in patches that are environmentally identical but an important synthetic point of metacommunity

theory in its general form is to highlight how heterogeneity of patches can alter these results.

- 7) These interactions affect which species can have self-sustaining populations and regulate the abundances of such populations in the absence of dispersal. Coexistence of populations of multiple species with such self-sustaining populations in local communities involves the interaction of stabilizing (frequency dependent) forces as well as fitness equalizing forces (Chesson 2000). *This expands and makes more precise Scheiner and Willig's 'trade-off principle' for diversity theory ((Scheiner & Willig 2005))*. In the presence of immigration some populations will also be present despite violating this proposition at the local scale.
- 8) In the extreme when stabilizing forces are weak and fitness equalizing forces are strong, stochastic demographic dynamics may affect patterns of distribution and abundance (Hubbell 2001). This derives in part from the third, fifth and sixth 'fundamental principles' of Scheiner and Willig (2005 (Scheiner & Willig 2005)).
- 9) Extinctions in local communities occur for at least three reasons. First, they may occur following environmental change or disturbance (Lande 1993), second they may occur as community composition changes via community assembly (Law & Morton 1996) and finally they may occur due to demographic stochasticity ((Lande 1993)). *This derives from some complex combination of the various 'fundamental principles' of Scheiner and Willig (2005) (Scheiner & Willig 2005) but is also strongly supported by basic theory of populations*. In the absence of dispersal, local communities consequently tend to lose species and become depauperate.

MODELS OF METACOMMUNITY THEORY

The propositions above can generate a rich array of possible outcomes. In many cases these differ substantially and sometime qualitatively from what is predicted in the absence of dispersal. But there is a rich array of outcomes that can result even within metacommunity theory depending on different assumptions and ranges of parameter values. To date there is no general model of metacommunity ecology; that is to say that there are a number of different mathematical formulations for metacommunity dynamics that make different assumptions but no mathematical uber-model that encompasses all of them. Instead the models available differ in

assumptions and in a very general way these models can be classified into four classes or archetypal types described in Table 8.1 (Leibold *et al.* 2004).

Models of patch dynamics

Here the focus is on colonization and extinction processes in patches. Dispersal is important because it allows recolonization following extinctions and extinctions are stochastic and due to disturbances or demographic stochasticity. The most well known examples include (Levins & Culver 1971; Hastings 1980; Tilman & Kareiva 1997). These models are structured in a way that completely ignores source-sink relations altogether so that they ignore proposition 2 above. Furthermore, for the most part, these models also ignore environmental differences among patches and assume that all patches are identical except for their occupancy. Such models can be modified to account for heterogeneity among patches (Holt 1993a; Chase & Leibold 2002) and such modifications can alter our understanding of some phenomena such as the prevalence of alternate stable states in metacommunities (Chase et al. 2002) but much more could be done along these lines. The most important insight to come from such models is that highly dispersive but poor competitors (also known as 'fugitive species') can coexist in a metacommunity with species that are superior competitors which would otherwise exclude (reviewed by (Mouquet et al. 2005)). In the case where there are environmental differences among patches, these models also predict that local community composition may not be perfectly associated with environmental differences since species may exist, albeit temporarily, in patches where they will eventually be displaced by superior competitors.

Models involving mass effects

Here the focus is on the role of dispersal in maintaining source-sink relations among populations in different patches. The emphasis is thus on propositions 2 and 4. Such mass effects allow for more species to coexist locally than might be predicted by theory on closed (no dispersal) communities and thus produce patterns that differ from those of community theory that ignores dispersal. On the other hand, as dispersal among such patches increases, the heterogeneities are effectively homogenized as far as they affect the overall mean fitness of populations across the entire metacommunity and eventually these heterogeneities are not sufficient to allow niche partitioning by patch type. To date these models have largely ignored stochastic extinctions due to disturbances and environmental change and they thus tend to ignore proposition 1 above. As in models of patch dynamics there may also be imperfect correspondence between local environmental conditions and community composition since deviations from these would result due to the existence of sink populations.

Models of species sorting in metacommunities

A third set of models conforms to the most conventional perspective about local communities by examining how species composition of communities respond to environmental gradients and heterogeneities. The idea is to examine how community assembly occurs in the absence of source-sink relations and in the absence of background extinctions. Here dispersal is only important because it provides the stream of potential colonists that allows community composition to track environmental changes in time and space. These models thus ignore propositions 2-4 above and focus on the range of parameters that prevent dispersal from being a limiting process in community assembly. The prediction is that local community composition should strongly track local environmental conditions. While this may appear to make this approach unlikely, it corresponds most strongly to conventional ideas about species distributions in relations to environment. Further, this approach is the one that is most often strongly supported by observations of the relation between community composition and environment ((Cottenie & De Meester 2005)).

Neutral metacommunity models

A final and most recent approach to metacommunities has been developed by Hubbell (2001). It focuses on the case when all species in a metacommunity have identical dispersal, extinction and competitive abilities. It thus ignores or re-interprets most of the propositions listed above. Dispersal still acts to fuel the arrival of species into local patches (proposition 1) and migrating individuals can allow species to exist that would otherwise go extinct (somewhat like proposition 2 but without fitness differences between patches), but the important factors that regulate community composition are chance demographic events and chance dispersal events. In strong contrast to species sorting models, there should be little correspondence between community composition and local environment since all species see environmental heterogeneities in identical ways. Again, while this may appear to make this approach unlikely, data on relative

abundance distributions show surprising fit to the predictions of this model (although other models can also show strong, and perhaps sometimes stronger, fit to the data).

Comparing the four archetypal models

The differences among these archetypal models thus depend on the role hypothesized for dispersal and the importance given to heterogeneity in local environmental conditions among communities. A very crude conceptual landscape for these models that illustrates possible relationships among these models is shown in Figure 8.2. Here I have identified two axes, connectivity (or inversely, dispersal) and environmental heterogeneity as characterizing the relevance of these different models.

Connectivity and dispersal play an important role in separating patch dynamic and species sorting from source-sink models when dispersal is either much less than local population turnover rates versus comparable or greater than this turnover rate. This is because it is only when immigrating rates are comparable with turnover rates that sink-populations are regular features of local communities. Similarly, dispersal plays an important role in separating patch dynamics from species sorting and mass effects when colonization due to dispersal is either comparable or much greater than local extinction rates. This is because colonization rates that greatly exceed local extinction rates likely constrain the possibility of 'fugitive' species and minimize the rate of community assembly (i.e. succession) to end-state configurations (e.g. 'climax communities').

Environmental heterogeneity plays an important role in separating neutral models and most patch dynamics models from species sorting and source-sink models. It is important to understand that what we mean by environmental heterogeneity is the heterogeneity that differentially affects the fitness of different species. Environmental heterogeneities may quantitatively affect species in the metacommunity but would not lead to species sorting or source-sink effects is the effects are symmetrical for all the component species.

A third conceptual axis that is not shown in Figure 8.2 but also important to discuss is whether the models are explicitly spatially structured or whether spatial effects are implicit. To date, much of the attention given to metacommunity models treat space implicitly. They most often do this by hypothesizing that each local community contributes to a single pool of dispersers that then disperses itself back into the local communities with equal probability. A more realistic view would model dispersal as being more likely among local communities that are more highly connected. A substantial body of work in spatial ecology exists that illustrates these effects. In the case of models without environmental heterogeneity among patches some general results can be obtained that are more precise than those obtained by the implicit models but this is more difficult to do with environmentally heterogeneous landscapes since now the precise arrangement of sites relative to each other can matter greatly. In many cases, however there is a general correspondence, at least in qualitative terms, between the implicit and explicit models. One exception are effects that depend on spatial patterning per se, that arise in explicit models but that simply cannot be found in implicit ones. An example of this can be seen in a simple model of non-transitive competition among three species (i.e. the 'rock-scissors- paper' game) in a spatial setting ((Kerr et al. 2002)) where long term coexistence of the three species depends on patterned 'traveling waves' of each of the species sequentially replacing each other in local communities. In an implicit model, these traveling waves do not exist and instead the overall frequency of species in the metacommunity oscillates in heteroclinic cycles. The amplitude of these cycles is not stable so that eventually only one of the species is present in the metacommunity ((Law & Leibold 2005))

Clearly there is a great need to better understand the ways these various archetypal models relate to each other and to try to synthesize them in various ways. Some progress has been made along these lines but much more is needed. Perhaps the best synthesis has been at the interface of species sorting and source-sink models ((Amarasekare & Nisbet 2001; Mouquet & Loreau 2002, 2003)),. I use this interface to illustrate how theory, models and experiments has led to improved understanding of diversity patterns in metacommunities next.

ILLUSTRATING LINKS BETWEEN THEORY, MODELS AND EXPERIMENTS

Metacommunity theory has elucidated a number of phenomena in community ecology (see (Holyoak *et al.* 2005) for a partial review). To illustrate how it can do so, I focus on the synthesis of the species sorting and mass effects models to explain how diversity varies with dispersal (or inversely, connectance) developed by Mouquet, 2003 #127; Mouquet, 2002 #142. This work combined analytical and numerical methods to explore how metacommunity structure varied as a function of dispersal among local communities. A few simplifying assumptions were made including:

- a) all species compete on a one for one basis
- b) all species have identical dispersal rates and all patches are similarly connected to the dispersal pool.
- c) each species has one patch type in which it is the best competitor. There is no niche partitioning within patches.
- d) initially, each species was the sole occupant of the patch in which it was dominant
- e) there is no environmental change within patches and thus no extinctions due to processes other than competition.

This synthetic model showed that species sorting and source-sink models are closely linked models that can explain how different levels of diversity vary with dispersal (Figure 8.3). In the absence of dispersal local diversity is low (one species per patch) and regional diversity is high (as many species as there are patch types). At low dispersal this is still true because immigrants remain rare and ephemeral in local patches. Local diversity increases when percapita dispersal rates are comparable to per-capita turnover rates. This is because there are increasingly large sink populations maintained in local patches due to immigration from other patches (sources) by species that would otherwise be excluded by local dominants. At some point however, as dispersal increase even more, these sink populations become proportionately larger and they affect the fitness of local dominants. This is exacerbated by the fact that local dominants also disperse more and thus emigrate rather than contribute to sustenance of local populations. Ultimately, if dispersal is sufficiently high, the organisms involved respond to an average of environmental conditions across all patches rather so that individual patch effects are no longer important. Thus fitness differences across the patches are increasingly homogenized by dispersal. In the end the metacommunity is essentially equivalent to a single community because individuals show no bias towards being present in local patches where they are competitive dominants. This scale transition reduces regional diversity (and consequently local diversity) to a single species but regional diversity (sometimes called 'gamma diversity') declines over a large range of dispersal values. The model thus predicts that regional diversity is not affected by low levels of dispersal but that it begins to decline at some medium to high level of dispersal. Local diversity (also called 'alpha diversity') increases with dispersal up to the point where regional diversity begins to decline, whereupon it also declines with regional diversity. The average dissimilarity in composition among communities (also called 'beta

diversity' starts out high and decreases in opposite to regional diversity (Figure 8.3). It is important to understand that in this model dispersal cannot contribute to local community assembly (Proposition 6) since local community assembly is assumed to have reached an assembly end point at the beginning of the simulations so that this model can only explore the consequences of Propositions 7 and 8.

Several experiments have tested the predictions of this model on patterns of diversity (reviewed by Cadotte 2006a). Figure 8.4 shows the results of one of these experiments conducted with protists in microcosms ((Cadotte 2006b)). Cadotte created a metacommunity consisting of small microcosms that were interconnected by tubing. Valves in the tubing allowed the manipulation of dispersal by being open for different proportions of time. The microcosms were inoculated with a diverse set of protists in either a uniform fashion (all microcosms initially had all species) or differentially (each microcosm differed in initial composition. Dispersal was either absent, low (presumably lower than or comparable to the population turnover rate) or high. His results show qualitative correspondence with the predictions of Mouquet and Loreau (2002) even though it is likely that organisms varied in dispersal rate and there was not likely to be strong habitat specialization.

Metacommunity theory can make predictions about numerous other features of community variability including coexistence and similarity (e.g. Leibold 1998, Hubbell 2001), abundance distributions (e.g. Hubbell 2001, McGill et al. 2006), species richness (e.g. Hubbell 2001, Chase and Leibold 2003), etc. In many cases these predictions differ substantially from what would have been predicted by models of closed communities. They thus provide a rich array of novel ways to evaluate data and experimentally study ecological processes in communities.

SPATIAL VS ENVIRONMENTAL REGULATION OF COMMUNITY STRUCTURE IN METACOMMUNITIES

One of the more important ways that metacommunity theory has informed community ecology is by suggesting that community composition could variously depend on purely spatial effects as well as the more conventionally studied environmental effects ((Cottenie & De Meester 2005)). As outlined above, the four archetypal models predict that the relative magnitude of these two types of effects should vary among them. Species sorting predicts that only environmental variation among patches should determine variation in community composition whereas the neutral models predict that there should be no such environmental regulation and that spatial effects should predominate. Mass effects predict that both environmental and spatial effects should be present. Patch dynamics models that ignore environmental variation of course predict only spatial effects but as pointed out above, it is possible to model patch dynamics in an environmentally heterogeneous set of patches and under such conditions both spatial and environmental effects should be present.

A large number of studies have been conducted to evaluate the relative importance of spatial and environmental effects. These studies should be interpreted with some caution because results depend a lot on which patches are chosen for inclusion in the study, which species are chosen, which environmental factors are measured, and how spatial effects are quantified but they can still give insights into how we evaluate community regulation. Cottenie (Cottenie 2005) has conducted a meta-analysis of 158 such studies and, inspired by metacommunity theory, there have since been quite a few more. His findings indicate that environmental effects tend to be more important than spatial one in the majority of cases. However he found that spatial effects were also relatively common. The conclusion indicates that natural metacommunities vary in how they are regulated but do not help identify just how they do so very well. Current work is seeking to find more refined analyses that might improve on these simple methods.

LINKING METACOMMUNITY THEORY TO GENERAL ECOLOGICAL THEORY Obviously this depends on how we define the domain and propositions of Ecology Theory. It is obvious that the links with the definition of Scheiner and Willig are extremely tight. However Ecology can also be viewed a bit more broadly as the interactions between organisms and their environment. Under this perspective, the links are weaker but still strong since they focus to any of the propositions that link organism-environment interactions to their distribution in space and time.

LINKING METACOMMUNITY THEORY TO OTHER CONSTITUTIVE THEORIES Again there are obviously strong links with many of the other constitutive theories in ecology. Particularly strong ones exist among some of those described in this book.

- Niche Theory (Chase Chapter 5) and predator-prey interactions (Holt Chapter 7): Niche theory is concerned with describing how organisms maintain populations and how in doing so they alter the environments they live in. A critical way they do this is by interacting with their prey and their predators (Elton 1923). By and large, our current understanding of niche theory is primarily applicable at the local scale (e.g. Hutchinson 1959, Tilman 1982, Chase and Leibold 2003) and dispersal is much less often considered as part of the niche of an organisms. This is not quite true since dispersal has often been thought important in niche relations in successional processes (Watt 1947, Pickett et al. Chapter 9) and has more recently been invoked in models of niche relations involving 'spatial storage' effects (e.g. Chesson 2005). However, as I have described it in this chapter, metacommunity theory consists of exploring how niche theory applied at local scales interacts with dispersal in larger spatial settings.
- Population Dynamics (Hastings Chapter 6). Here the links are clearest as links between metapopulation theory and metacommunity theory involving patch dynamics and their domains and propositions are very similar. Indeed such metacommunity models of patch dynamics have to date been extensions of metapopulation models. One intriguing contribution that metacommunity approaches may make is to enrich the metapopulation models by focusing on the role of patch type heterogeneity. When the focus is on single species, heterogeneity in patch quality is likely to be less important than when multiple species are involved. Another contribution is the suggestion that a landscape of patches that may appear to one species as a metapopulation, may appear to another species with which it interacts as a much more homogenous single patch.
- Succession theory (Pickett et al. Chapter 9). As was initially argued by Watt (1947) the successional process implicitly involved spatio-temporal processes that depend on dispersal among localities with different local conditions. While the propositions of successional theory are strongly linked to metacommunity theory the domains differ somewhat in the degree of attention they give to temporal vs spatial dynamics.
- Island Biogeography (Sax and Gaines Chapter 10). Here the interest in dispersal is a strong linking aspect with metacommunity theory but the propositions are quite different. Island biogeography does not necessarily focus on local communities per se since it is more concerned with islands that may (and probably generally do) include multiple

habitats and patches with distinct sets of locally interacting species. Thus many of the complications and contingencies that are key propositions to at least some metacommunity models are ignored. Further, the role of dispersal in island biogeography is primarily the same as in models of patch dynamics but different than that in mass effects and species sorting models where stochastic local extinctions are absent.

- Global Change (Peters et al. Chapter 12). Peters et al. describe the role for dispersal and community change in global ecology in their chapter. The importance here is that responses of biotas (including physiological and genetic responses as well as community responses) and their consequences for ecosystem processes, are not locally constrained. Thus while the domain is different in emphasis on anthropogenic effects at the global level and on the consequences for ecosystem processes, the propositions are remarkably similar.
- Ecological Gradients (Fox et al. Chapter 13): Obviously, metacommunity models of species sorting and mass effects as well as patch dynamics models that include patch heterogeneity imply the existence of environmental gradients as expounded by Fox et al. (Chapter 13). The connection is particularly strong with gradient models that include species interactions. However the domains of the two theories are a bit different. Constitutive theories of ecological gradients (as described by Scheiner and Willig (2005) and Chapter 13 focus exclusively on correlations between species richness (only one aspect of community composition) and environment and do not address sources of variation in communities that are due to spatial effects. Thus the domain of metacommunity theory is a bit larger than that described in Chapter 13. Further, dispersal plays no key role in the constitutive theory of gradients described in Chapter 13 although as argued above it is probably important in fueling community assembly and species sorting.
- Biogeographic Gradient Theory (Colwell Chapter 14): Although metacommunity and biogeography theory seem to share many aspects and propositions, including dispersal, environmental contingency, spatial effects etc. there is a fundamental difference in the domain. Metacommunity theory as formulated in this chapter focuses on purely ecological effects and primarily examines their consequences in ahistorical settings. The spatial extent that such a focus implies ignores the phylogenetic and historical
biogeographic contingencies that characterize biogeographic gradient theory. An implied assumption is that all organisms have had equal likelihood of access to all sites within the metacommunity and this constrains the spatial extent of metacommunity theory to being at sufficiently small scales where this is reasonable. In contrast, biogeographic gradient theory is primarily concerned with larger spatial scales where phylogeny and historical effects are important. An interesting possibility is that these scales will overlap when a set of species that has a distribution over a particular area that is regulated by metacommunity processes interacts with another who's distribution in that same area is regulated by biogeographic gradients. Just such a case is suggested by recent analyses of zooplankton distributions in the Northwestern USA in which calanoid copepods have distributions that seem to reveal strong historical biogeographic effects and little environmental regulation whereas coexisting daphnid cladocerans show just the reverse (Leibold et al. in review).

A PROSPECTUS FOR METACOMMUNITY THEORY

The possible synthesis of metacommunity theory suggested by Figure 8.2 also points out that the biggest lacuna in our understanding of metacommunity theory is understanding what happens when different species differ strongly in dispersal rate. Several models and hypothesis that build on metacommunity theory show that variability in dispersal can be important in various ways:

- Colonization-competition relations: These have been hypothesized since the verbal models of (Diamond 1975) and the mathematical models of Hastings ((Hastings 1980)) based on colonization-extinction dynamics. More recent work has also suggested that such a trade-off might allow for coexistence in models based on birth-death-dispersal dynamics ((Amarasekare & Nisbet 2001; Mouquet & Loreau 2002),) but here the mechanism is totally different. Recent work in this area is showing that both of these types of effects can allow for coexistence in a metacommunity but that the conditions for doing so can be somewhat constrained (Yu & Wilson 2001) because it involves careful 'titration' of the relative rates of competitive exclusions and dispersal.
- Heterogeneous dispersal among trophic levels: It has been suggested that an important stabilizing feature of food webs emerges from differences in dispersal rates of different trophic levels (McCann *et al.* 2005). This issue is complicated because different effects

occur depending on whether dispersal is random or biased by adaptive or maladaptive behavior (Amarasekare 2008).

These findings illustrate a complex array of possibilities but it is difficult to yet see how they act in concert or how important these possibilities are in real communities. Future work on these questions will, no doubt, elucidate how dispersal interacts with other fundamental aspects of ecology but it is still hard to know just what this will involve.

Additionally, it is clear that metacommunity theory has many elements and parallels with evolutionary dynamics in metapopulations. Evolution in metapopulations also involves the role of dispersal in maintaining genetic variation on which selection can act and in homogenizing such variation in ways that roughly correspond to the roles of dispersal in metacommunities outlined above (Urban et al. 2008b). A suite of novel interactive dynamics between such adaptive dynamics and metacommunity processes emerge (e.g. Holt et al. 2005, Loeuille & Leibold 2008, Urban et al. 2008a). Two lines of evidence support the likelihood that there can be a strong interaction between evolutionary and ecological dynamics along these lines. The first is a growing body of work under the rubric of 'community genetics' (Whitham et al. 2003) that suggests that genotypic variation within some species can alter community properties associated with that species in ecosystems. The other is another growing body of work under the rubric of 'geographic mosaic evolution' (Benkman et al. 2001; Thompson 2005) that suggests that community context can strongly alter evolutionary and coevolutionary responses of species in communities over space. A few studies that explicitly explore the interface of metacommunity and evolutionary dynamics (De Meester et al. 2007) indicate that there can be important interactions between these processes. If this is substantiated by further work, additional propositions about the role of adaptive evolution on metacommunities will need to be explored.

Finally, a challenge for metacommunity theory emerges from its focus on larger spatial scales. Expanding from the local to metacommunity scales adds a number of important ideas about how purely ecological processes may explain variation in the distribution of organisms among communities in an area. However such purely ecological explanations will only apply up to a certain spatial scale; eventually other processes involving biogeographical and phylogenetic dynamics will also play a role and at some point will overwhelm any explanations that derive from metacommunity ecology. More intriguing is the likelihood that different groups of

interacting organisms in metacommunities might differ in the spatial scale where this happens. Leibold and Economo (in review) for example, found that distributions of daphniid zooplankters were strongly related to environmental conditions as predicted by 'species sorting' models of metacommunities in the Northeastern US whereas calanoid copepod zooplanktors in the same lakes showed distributions that were strongly constrained by biogeographic dynamics from post-Pleistocene refaunation. If there are strong interactions between daphniids and calanoids this would suggest that some metacommunity dynamics might be shaped by biogeographic propositions as well.

Metacommunity theory is relatively new in ecology. In this chapter I have tried to identify the elements of a constitutive theory of this topic but there is to date no comprehensive model that integrates the main propositions. Nevertheless what work there is highlights the insights that can come from metacommunity theory and illustrate the potential for a deeper understanding of meso-scale ecology as this theory develops in the future. Literature Cited

- Amarasekare P. (2008). Spatial Dynamics of Foodwebs. *Annual Review of Ecology, Evolution, and Systematics*, 39, 479-500.
- Amarasekare P. & Nisbet R.M. (2001). Spatial heterogeneity, source-sink dynamics, and the local coexistence of competing species. *American Naturalist*, 158, 572-584.
- Beisner B.E., Peres P.R., Lindstrom E.S., Barnett A. & Longhi M.L. (2006). The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology*, 87, 2985-2991.
- Benkman C.W., Holimon W.C. & Smith J.W. (2001). The influence of a competitor on the geographic mosaic of coevolution between crossbills and lodgepole pine. *Evolution*, 55, 282-294.
- Cadotte M.W. (2006a). Dispersal and Species Diversity: A Meta-Analysis. *American Naturalist*, 167, 913-924.
- Cadotte M.W. (2006b). METACOMMUNITY INFLUENCES ON COMMUNITY RICHNESS AT MULTIPLE SPATIAL SCALES: A MICROCOSM EXPERIMENT. *Ecology*, 87, 1008-1016.
- Chase J.M., Abrams P.A., Grover J.P., Diehl S., Chesson P., Holt R.D., Richards S.A., Nisbet R.M. & Case T.J. (2002). The interaction between predation and competition: a review and synthesis. *Ecology Letters*, 5, 302-315.
- Chase J.M., and Leibold, M.A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago.
- Chase J.M. & Leibold M.A. (2002). Spatial scale dictates the productivity-biodiversity relationship. *Nature*, 416, 427-430.
- Chesson P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31, 343-+.
- Cornell H.V. & Lawton J.H. (1992). SPECIES INTERACTIONS, LOCAL AND REGIONAL PROCESSES, AND LIMITS TO THE RICHNESS OF ECOLOGICAL COMMUNITIES -A THEORETICAL PERSPECTIVE. *Journal of Animal Ecology*, 61, 1-12.
- Cottenie K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8, 1175-1182.

- Cottenie K. & De Meester L. (2005). Local interactions and local dispersal in a zooplankton metacommunity. *Metacommunities: Spatial Dynamics and Ecological Communities*, 189-211.
- De Meester L., Louette G., Duvivier C., Van Darnme C. & Michels E. (2007). Genetic composition of resident populations influences establishment success of immigrant species. *Oecologia*, 153, 431-440.
- Diamond J. (1975). Assembly of species communities. In: *Ecology and Evolution of Communities* (ed. JM CMaD). Harvard University Press Cambridge.
- Hanski I.a.G., Michael (1991). Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society*, 42, 3-16.
- Hastings A. (1980). DISTURBANCE, COEXISTENCE, HISTORY, AND COMPETITION FOR SPACE. *Theoretical Population Biology*, 18, 363-373.
- Holt R.D. (1984). SPATIAL HETEROGENEITY, INDIRECT INTERACTIONS, AND THE COEXISTENCE OF PREY SPECIES. *American Naturalist*, 124, 377-406.
- Holt R.D. (1993a). Ecology at the mesoscale: the influence of regional processes on local communities. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, 101, 77–88.
- Holt R.D. (1993b). Ecology at the mesoscale: The influence of regional processes on local communities. In: Species Diversity in Ecological Communities: Historical and Geographical Perspectives (ed. Ricklefs REaSD). University of Chicago Press Chicago, II.
- Holt R.D., Holyoak M. & Leibold M.A. (2005). Future directions in metacommunity ecology. *Metacommunities: Spatial Dynamics and Ecological Communities*, 465-489.
- Holyoak M., Leibold M.A. & Holt R.D. (2005). *Metacommunities: Spatial Dynamics and Ecological Communities*. University Of Chicago Press.
- Hubbell S.P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton.
- Kerr B., Riley M.A., Feldman M.W. & Bohannan B.J.M. (2002). Local dispersal promotes biodiversity in a real-life game of rock--paper--scissors. *Nature*, 418, 171-174.
- Kneitel J.M. & Miller T.E. (2003). Dispersal rates affect species composition in metacommunities of Sarracenia purpurea inquilines. *American Naturalist*, 162, 165-171.

- Lande R. (1993). RISKS OF POPULATION EXTINCTION FROM DEMOGRAPHIC AND ENVIRONMENTAL STOCHASTICITY AND RANDOM CATASTROPHES. American Naturalist, 142, 911-927.
- Law R. & Leibold M.A. (2005). Assembly dynamics in metacommunities. *Metacommunities: Spatial Dynamics and Ecological Communities*, 263-278.
- Law R. & Morton R.D. (1996). Permanence and the assembly of ecological communities. *Ecology*, 77, 762-775.
- Leibold M.A., Holyoak M., Mouquet N., Amarasekare P., Chase J.M., Hoopes M.F., Holt R.D., Shurin J.B., Law R., Tilman D., Loreau M. & Gonzalez A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7, 601-613.
- Levins R. (1975). Evolution in communities near equilibrium. In: *Ecology and Evolution of Communities* (ed. Cody MaDJ). Belknap Press of Harvard University Press Cambridge, pp. 16-50.
- Levins R. & Culver D. (1971). Regional Coexistence of Species and Competition between Rare Species. *Proceedings of the National Academy of Sciences*, 68, 1246-1248.
- Loeuille N. & Leibold M.A. (2008). Evolution in metacommunities: On the relative importance of species sorting and monopolization in structuring communities. *American Naturalist*, 171, 788-799.
- McCann K.S., Rasmussen J.B. & Umbanhowar J. (2005). The dynamics of spatially coupled food webs. *Ecology Letters*, 8, 513-523.
- Mouquet N., Hoopes M.F. & Amarasekare P. (2005). The world is patchy and heterogeneous! Trade-off and source-sink dynamics in competitive metacommunities. *Metacommunities: Spatial Dynamics and Ecological Communities*, 237-262.
- Mouquet N. & Loreau M. (2002). Coexistence in metacommunities: The regional similarity hypothesis. *American Naturalist*, 159, 420-426.
- Mouquet N. & Loreau M. (2003). Community patterns in source-sink metacommunities. *American Naturalist*, 162, 544-557.
- Ricklefs R.a.S.D. (1993). Species Diversity in Ecological Communities: Historical and Geographical Perspectives. University of Chicago Press, Chicago.
- Scheiner S.M. & Willig M.R. (2005). Developing unified theories in ecology as exemplified with diversity gradients. *American Naturalist*, 166, 458-469.

Skellam J. (1951). Random dispersal in theoretical populations. *Biometrika*, 38, 196-218.

- Snyder R.E. & Chesson P. (2004). How the Spatial Scales of Dispersal, Competition, and Environmental Heterogeneity Interact to Affect Coexistence. *American Naturalist*, 164, 633-650.
- Thompson J.N. (2005). Geographic Mosaic of Coevolution. In: *Geographic Mosaic of Coevolution*.
- Tilman D. & Kareiva P.M. (1997). Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions. Princeton University Press.
- Urban M.C., Leibold M.A., Amarasekare P., De Meester L., Gomulkiewicz R., Hochberg M.E., Klausmeier C.A., Loeuille N., de Mazancourt C. & Norberg J. (2008a). The evolutionary ecology of metacommunities. *Trends in Ecology & Evolution*.
- Urban M.C., Leibold M.A., Amarasekare P., De Meester L., Gomulkiewicz R., Hochberg M.E.,
 Klausmeier C.A., Loeuille N., de Mazancourt C., Norberg J., Pantel J.H., Strauss S.Y.,
 Vellend M. & Wade M.J. (2008b). The evolutionary ecology of metacommunities. *Trends Ecol Evol*, 23, 311-7.
- Whitham T.G., Young W.P., Martinsen G.D., Gehring C.A., Schweitzer J.A., Shuster S.M.,
 Wimp G.M., Fischer D.G., Bailey J.K., Lindroth R.L., Woolbright S. & Kuske C.R. (2003).
 Community and ecosystem genetics: A consequence of the extended phenotype. *Ecology*, 84, 559-573.
- Yu D.W. & Wilson H.B. (2001). The competition-colonization trade-off is dead; Long live the competition-colonization trade-off. *American Naturalist*, 158, 49-63.

Table 8.1. A summary of the four paradigms for metacommunity theory (based on Leibold et al. 2004). The Propositions listed correspond to those listed in the text. Cases marked with a question mark (?) indicate that the answer depends on whether patch dynamic models involve homogenous patches (the most commonly studied case) or patches that vary in environmental conditions (also possible). Predictions of each model type are selected ones meant to highlight differences among the models. Actual metacommunities are likely to consist of mixtures of all of these paradigms so that some groups of species may have distributions that are more strongly related to different models.

Paradigm	Patch dynamics	Mass effects	Species sorting	Neutral
1) dispersal affects colonists	Х		Х	Х
2) Dispersal allows source-sink		Х		Х
relations				
3) Heterogeneity in dispersal is	Х	Х		
important				
4) dispersal allows spatial	Х	Х		
feedbacks				
5) Interactions are direct and	Х	Х	Х	
indirect				
6) Interactions depend on local	?	Х	Х	
environments				
7) coexistence requires stabilizing	Х		Х	
effects in local communities				
8) stochastic demography is				Х
important in allowing				
coexistence				
9) local extinctions are important	Х			

Predictions:

Invasible local communities should	Y	Ν	Ν	Y
be common				
Fugitive species should exist in the	Y	Ν	Ν	Ν
metacommunity				
Composition of communities	Y	Y	Ν	Y
should depend on spatial effects				
independent of environment				
Composition of communities	?	Y	Y	Ν
should depend on				
environmental effects				
independent of spatial effects				
The ratio of local to regional	Ν	Y	Ν	Y
diversity can be high				



Figure 8.1 Use of the concept 'metacommunity' through time as cited in Web of Science.



Connectivity

Figure 8.2. The four major 'paradigms' of metacommunity ecology in relation to the amount of interpatch habitat heterogeneity and connectivity. Patch heterogeneity refers to the degree of interpatch differences that differentially favor different species in different patches. Connectivity can be quantified as the rate of immigration (events/time) and can be compared to two other rates, the local extinction rate (E) and the local death rate (D). PD = patch dynamic, SS = species sorting, ME = mass effects, NM = neutral model. A separate oval has been drawn for PD models that focus on competition-colonization models in the absence of patch heterogeneity to highlight the fact that most PD models are of this type. The oval for SS has purposefully been drawn so as to extend to the right of connectivity equal to D and overlap with ME to highlight that high dispersal can still produce patterns and processes similar to SS if dispersal is adaptive. See Table 8.1 for a description of the assumptions and critical processes that apply to each 'paradigm'.



Figure 8.3. The hypothesized interaction between dispersal rate and species diversity at different spatial scales (taken from (Cadotte 2006a)). Local diversity here is the same as ' α -diversity' in the text, and regional diversity is the same as γ -diversity in the text. The dispersal rate here is conceptually identical with 'connectivity' as used in Figure 2. The downturn in Beta diversity and the upturn in Local diversity in the simulations that generated this graph correspond to the per capita turnover rate of individuals in the populations.



Figure 8.4 Experimental results of manipulating dispersal in a protist microcosm (Cadotte 2006). Dispersal was manipulated by connecting microcosms with tubing and was either absent (Control), low (clamps on the tubing were used to control dispersal at 1 hour every two days) or high (tubing was never closed). Initial conditions either had all subpopulations started with identical communities (Initial Beta = 0) or they differed by restricting species inoculations to different microcosms with an initial Beta = 6. Multiway ANOVA indicated that data points labeled with similar letters are not statistically different.

Chapter 9: Domain and Propositions of Succession Theory Steward T. A. Pickett, Scott J. Meiners, and Mary L. Cadenasso

INTRODUCTION

Succession is perhaps the oldest of ecological concepts, having arisen when ecology was emerging as a self-conscious discipline (McIntosh 1985). Yet, it continues to address many fundamental issues in ecology, to support important applications, and to synthesize the insights and perspectives of other theories. Thus, it fulfills two functions key in assessing the utility of a contemporary ecological theory. First, it exhibits the attributes of a mature, well developed, and intensively tested theory (Glenn-Lewin et al. 1992, Pickett and Cadenasso 2005). Second, it provides a linkage among theories and applications that have usually been considered separately (Walker et al. 2007). For example, the theory of succession or community dynamics has been applied to terrestrial and aquatic habitats (Bazzaz 1979, Stevenson et al. 1991, Biox et al. 2004), and for microbial (Boucher et al. 2005), bird (Keller et al. 2003), soil invertebrate (Yi et al. 2006), and mammal (Schweiger et al. 2000) communities.

This chapter outlines the structure of contemporary succession theory, beginning with a statement of its domain. It proceeds through an elaboration of the individual propositions of the theory. The term, proposition, is a general one that encompasses both the conceptual and the hybrid conceptual-empirical components of theory. Based on a broad view of theory (Pickett et al. 2007), propositions may be assumptions of domain, statements of concept, definitions, empirical generalizations, laws, and hypotheses or expectations derived from models. The definition of proposition used in this book is narrower (Scheiner and Willig, Chapter 1, this volume), so we will address domain as a background assumption. The propositions we present are simply the broad principles that embody the key aspects of knowledge about community dynamics. The propositions can be divided among those that 1) identify the basic actors in succession; 2) enumerate the higher level environmental and contextual causes; and 3) describe the nature of successional trajectories and the outcomes of detailed causes. The propositions address issues that ecologists have struggled with and refined over decades. However, the contemporary content and assumptions embodied in the propositions differ substantially from those of earlier eras in ecology (Johnson 1979). Presenting propositions of a theory separately is an artificial device for rhetorical purposes. In fact, the propositions of a theory are closely linked to each other, and the meaning of each proposition is only fully realized when it is related to the other components of the theory (Pickett et al. 2007). To illustrate how this theory has developed, we also identify key differences between the propositions of classical succession theory as often presented in textbooks (Pickett and Cadenasso 2005, Eliot 2007) and contemporary versions of community dynamics theory (Table 9.1).

DOMAIN OF SUCCESSION THEORY

Bounding a discipline or a establishing a domain is the first job of a constituent theory. The domain statement of succession theory is a neutral definition of the concept. Neutral definitions are intended to identify the core meaning of a concept, minimizing assumptions about outcome or dominant mechanism (Jax et al. 1998). Following upon this definition of domain, it will become clear that succession is an extraordinarily broad concept, which has great synthetic power for ecology. Although most of our examples will come from plant communities, with which we are most familiar, it is clear that the propositions, when couched in terms that are organism independent, apply to any aggregation of species (Morin 1999).

Domain. Succession is the change in structure or composition of a group of organisms of different species at a site through time.

Succession or community dynamics is preeminently about the behavior of collections of species through time and can involve changes in species composition or the three dimensional structure, that is the architecture or physiognomy, of species assemblages. Often succession has been defined only as change in species composition (Begon et al. 1996) regardless of whether it is plants or animals that are the focus. However the founding concern with change in dominance of plant growth form makes it clear that the architecture of a species assemblage can be of equal importance for succession (Clements 1916). Indeed, changes in plant species composition in well watered habitats often result in dramatic architectural changes as the community shifts from herbaceous, to shrub, to tree dominance. Even in cases where one species is dominant, growth in stature, clonal growth, or canopy coalescence as individual woody plants mature are important successional changes (Muller 1952). Architectural complexity can characterize intertidal community succession after disturbance (Dayton 1975). Architectural changes such as these

have implications for resource availability, niche construction for consumer species, and susceptibility to subsequent disturbances (e.g. Pickett et al. 2001).

The definition of succession or community dynamics used here is a neutral one that focuses on the kernel of the process, and does not require the assumptions often attached to the definition that describe specific temporal patterns to obtain. The domain of succession theory based on the neutral definition used here, makes no judgment about progress, directionality, temporal scale, end point, or whether the species form a tightly integrated community or not. This approach focuses on both compositional and architectural changes in communities in the broadest, most generalizable way possible. This means that whether succession is progressive and directional, or takes place over temporal scales of ten to 10,000 years, and so on, are decisions about the structure of specific models and not the general theory. Indeed, as we discuss later, the general mechanisms of community change are independent of temporal scale.

We suggest that the persistent controversies about directionality and the existence of an identifiable endpoint that have plagued the definition of succession, be left as open issues to be dealt with by the explicit assumptions embodied in individual models. It is not succession in the broad sense that is problematic, but rather the contradictory assumptions of different models and applications that generate problematic controversy. Distinguishing between the core, neutral concepts and the specific models that translate them to real or simulated situations prevents the repeated discovery that classical "succession theory" doesn't work (McIntosh 1980), and hence should be discarded. In contrast, differentiation between the core concept, which is neutral and broadly applicable, and the more narrow specific models (Jax et al. 1998), permits appropriate application to an impressive range of real world situations, while allowing general similarities of process to be recognized (Pickett and Cadenasso 2005).

PROPOSITIONS ABOUT THE ACTORS IN THE PROCESS

The next set of propositions identifies the nature of the fundamental units and the kinds of interactions between units through succession. These propositions identify an inherently hierarchical arrangement, with the fundamental units or actors interacting on one level, and the results emerging on a higher level of organization. Community dynamics theory must account for processes, interactions, and constraints that exist in linked hierarchical levels: organismal

adaptation, physiology, and plasticity; population structure and density; community composition and architecture; landscape structure and fluxes; and ecosystem feedbacks and processes.

Proposition 1: Succession is driven by the interactions of individual organisms with each other and the physical environment.

The interaction of individuals is made manifest by the change in species composition observed in many successions. However, architectural changes not accompanied by change in composition also reflect interaction of individuals. An example would be a case of vegetation change where the stature and density of a monospecific stand changes through time. As the individuals in a collection grow, they occupy increasing space, until some are overtopped or otherwise disadvantaged in their interactions within the community. This is seen in the stand thinning that characterizes dense, young forests. During stand thinning, the height of the upper layer of the canopy, the depth of the canopy volume, the spread of surviving crowns, and the presence of downed debris and depth of soil organic matter may all increase. This is a structural or architectural succession with a constant composition. Bird and insect diversity, productivity, and decomposer communities are aspects of the ecosystem that may change along with the architectural changes occurring in the plant community (Odum 1969). Indeed, many of the expectations proposed by Odum (1969) can be explained by the growth of individuals, the accumulation of species with longer life spans, and the competition for limiting resources (Loreau 1998).

When succession was first conceived, much of the attention of ecologists was on the coarse scales of pattern in vegetation (McIntosh 1985). Ecology was still close to its roots in biogeography and was attempting to apply the familiar knowledge of plant physiology to explain those large global and continental patterns (Kingsland 2005). In light of this, inventing the concept of succession was a clear advance, because it exposed the dynamism that existed in vegetation at smaller scales. However, the predominance of the coarse scale focus in vegetation description led ecologists originally to articulate the causes of succession at the same coarse spatial scales upon which the changes were observed. Hence, they focused on communities. Changes in a community were presumed to be due to features of the community itself and the goal-seeking tendencies of those communities (McIntosh 1980). Hierarchy theory has since clarified the error of this approach as well as presented an alternative (Ahl and Allen 1996).

Successional explanations have been vastly improved by using a hierarchical strategy of explanation (Pickett et al. 1987a), in which interactions exist at the level of individual organisms, and their direct and indirect effects emerge as the higher level of community or ecosystem patterns. There can also be "downward" influences from larger scale ecosystem and landscape conditions to the community, and from community to population and individual in succession.

Although the earliest efforts to clarify the individual-based nature of vegetation dynamics (Gleason 1917) met with resistance, subsequent fine scale, long-term, and experimental research confirmed that interactions among individual organisms as parts of populations was the core process in succession (Horn 1974; Miles 1979; Parker 2004). This is called the individualistic approach to succession. Both Clements (e.g. 1916) and Gleason (e.g. 1917), according to a rigorous philosophical analysis, based their explanations of successional causes on the adaptive physiology of individual organisms (Eliot 2007). Indeed, they both recognized migration and establishment, and environmental sorting as key processes, as will be discussed later. According to Eliot (2007: 104) the difference is subtle: Gleason and Clements "differ in emphasis, in their prioritization of causal factors. While Clements emphasizes *environmental* sorting among potential individual immigrants, Gleason emphasizes environmental sorting among potential immigrants" (emphasis in original). Furthermore, they took different scaling approaches, with Gleason emphasizing local variation, and Clements structures his arguments from adaptation to large scale vegetation. Eliot's (2007) paper is an important advance in recognizing the similarities and differences between the approaches of these two pioneers, which are usually represented as combative caricatures in contemporary literature. The process based approach presented in the propositions below combines key elements from both the Clementsian and Gleasonian traditions (Pickett and Cadenasso 2005).

It is important in considering the individualistic basis of succession to recognize that it is the interactions, not merely the existence of individuals as separate entities that is of concern. Indeed, individuals in communities cannot be treated as independent atoms, because their density and frequency are important aspects of how they interact (e.g. Morin 1999). Likewise, the age or stage structure of the populations can be important in their interactions. The effects of one population on another can change with density and frequency, and the potential for indirect effects through third parties is also a part of the interaction environment of individuals. **Proposition 2**: Successional patterns in communities result from the interactions of individuals.

Proposition 1, laid out above, explicitly focuses on individuals and their interactions, while Proposition 2 draws the implications of the individual focus of that proposition for a different realm of ecological aggregation, that of the community. The implication of the individualistic proposition (No. 1), with its critique of the strictly community level explanation of succession, is that community changes must result from the individualistic behavior of organisms. A point of confusion is sometimes still expressed concerning the individualistic nature of species central to the proposition. Individuality does not mean that species are necessarily acting randomly or independently, although the neutral model raises that possibility (Hubbell 2001). In fact, it is the variation in interaction among species, their direct and indirect influences on one another and on their physical and chemical environment, that embody the most general mechanism of succession (Glenn-Lewin 1980; Huston 1992; Myster & Pickett 1988). Indeed, the concept of the ecosystem was invented to acknowledge the interaction of the biotic complex and the physical complex in the context of succession (Tansley 1935). Contemporary ecology also recognizes that levels of organization larger than individual communities, such as metacommunities, landscapes, or patch mosaics, can also contribute to community dynamics processes (Hansson and Angelstam 1991; Veblen 1992, Sousa 1984). This point is emphasized in later propositions.

Proposition 3: Multiple trophic levels participate in the driving interactions.

The concept of succession was invented by plant ecologists, and has been extensively used by ecosystem ecologists as well (Odum 1969). This sociological history seems to have led to an inappropriate neglect of higher trophic levels as active participants in many successions. Some early studies indicated that animal assemblages followed the compositional and structural changes wrought by the plant community, but they neglected the influences of animals in shaping succession (Smith 1928). Research in the last few decades has corrected this bias and oversight (Davidson 1993; Bowers 1997; Inouye et al. 1994; Meiners et al. 2000). The role of dispersers was perhaps the first to be recognized as important in directing succession. However, seed and seedling predators, herbivores including browsers and foliage consumers, and animals as ecosystem engineers are now widely accepted as potential agents of successional change. The potential for pathogens and mutualists to generate feedbacks affecting community dynamics is becoming more apparent (Reynolds et al. 2003; Bever 2003; Klironomos 2002; Dobson and Crawley 1994).

The incorporation of consumer trophic levels has, like the recognition of feedbacks with the physical environment, added significant dimensionality to the concept of succession. Direct and indirect effects via feedbacks with consumers, mutualists, and engineers mean that interactions are no longer conceptually restricted to the plants which so conspicuously structure many successions.

These two propositions reflect an individual-based or population approach, recognizing that multi-species assemblages of organisms of various trophic levels and including plants, animals, and microbes, generate any observed community changes (Johnson 1979). The organismal level of organization upon which succession is centered is itself a complex of interactions and feedbacks.

PROPOSITIONS ABOUT THE ENVIRONMENT AND CONTEXT AS CAUSES

The propositions to this point set the stage for understanding how succession occurs by identifying the individuals as the core actors in a linked hierarchy of biological entities. The next two propositions indicate the nature of interactions among the actors and between actors and environment that can develop during succession.

Proposition 4: Succession results from processes of disturbance, differential availability of species to a site, and differential performance of species within a site.

This proposition summarizes the core of the community dynamics by identifying the three most general causes of the process. Disturbances are particular events that alter the physical structure of a community or cause mortality of structurally or functionally dominant organisms (White and Jentsch 2001). Such events generate opportunities for the release of suppressed residents, or the establishment of new individuals, and hence, new species. It is the behavior of specific events relative to the structure of a community, and to the requirements of species that are available to respond to it, which determines what actually disrupts a particular community, and whether particular species are favored or disfavored. Disturbance as a concept is a useful generalization, but it is the characteristics and responses to particular events that contribute to the understanding of succession. The general process of disturbance recognizes the

particularities required through its hierarchical structure (Figure 9.1). It has nested within it the resource base of the site, the identity of the agents of disturbance, and the intensity, size, and timing of the potentially disturbing events (Pickett et al. 1989). Differentials in disturbance influence succession by affecting contrasting site features, as well as features of the disturbances themselves. The presentation of disturbance in succession illustrates a process-based approach. It suggests a hierarchy of general causes which can be broken down into more specific mechanisms, constraints, and enablers among which interactions may occur (Pickett and Cadenasso 2005; Pickett et al. 1987a).

The second process of community dynamics is differential availability of species at a site. This differential is driven by variation in species' inherent capacities for dispersal in space, dormancy through time, and proximity and landscape connectivity of sources of migrants to the study site (Hobbs and Walker 2007; Meiners et al. 2003; Pickett and Cadenasso 2005). This is as true of animals as it is of plants. Landscape configuration not only determines the distance over which dispersal is required, but may also generate resistance to dispersal across boundaries (Cadenasso et al. 2003). In addition, the identity and behavior of both biotic and abiotic dispersal vectors must be recognized as a part of this process (Figure 9.1).

The third process of community dynamics is differential species performance. This process aggregates the individual characteristics into populations, and acknowledges that population features, such as density, stages or ages, size classes, and frequency can influence the outcome of interactions. Differential performance encompasses all the processes of establishment, growth, survival, reproduction, and mortality that plants, animals, and microbes employ (Davidson 1993; Keever 1979; MacMahon 1980; Vitousek 2004). In includes interactions that are competitive, consumptive, and mutualistic, and the strategic, physiological, and life history foundations of these diverse interactive capacities. In addition, it includes adaptive and plastic potentials of species, and it recognizes the potential for stresses and fine scale disturbances to interact with species performance. Variation in one or more of these processes among an assemblage can result in change in the composition or structure of the community. Any factor that alters the performance attributes of one or more species can be incorporated into a more specific model of a particular succession; we do not attempt to enumerate all the possible specific causes in this chapter. Although disturbance acts before the other two causes and sets the stage for the change within the system, either differential

availability or differential performance of species is sufficient to generate successional change (Figure 9.2).

This outline of the three causes of community dynamics is, on one hand, an empirical generalization extracted from explanations of plant, animal, and microbial successions accumulated over more than a century of study (Glenn-Lewin et al. 1992; MacMahon 1981; Miles 1979). On the other hand, it is a recasting of the general causes of succession that Clements (1916) and Gleason (1917) recognized within two decades of the scientific framing of the concept. The fact that the triad of successional causes satisfies both empirical and conceptual adequacy for explanation argues for its generality. We hypothesize that no specific mechanism of succession identified in the future will fall outside this causal framework, but will merely add detail to the hierarchical list of causes. We expect this to apply to animal and plant communities, and to aquatic and terrestrial environments (Table 9.2).

This causal proposition also reflects the inadequacy of explanations based on only one kind of cause (Connell et al. 1987). Such one-factor explanations were encouraged by the simplistic, hypothetico-deductive philosophy of science in favor during the middle decades of succession study (Pickett et al. 1994). However, multicausal modeling and experimentation is now exposing the complexity of successional drivers (Clark 2007; Wyckoff and Clark 2002). The hierarchical framework, in which the general causes or processes of succession are disaggregated into more specific mechanisms and drivers to effectively construct working models, supports the multicausal approach.

Proposition 5: Successional causes can operate on any time scale.

Succession has traditionally been construed to act over time scales on the order of 1 to 10^3 years, while seasonal and paleoecological time scales have been excluded. However, we see no compelling reason for either exclusion. The causes of community dynamics are common to community dynamics on all these time scales (Brand and Parker 1995). Biogeographic patterns, post-glacial vegetation shifts, and alterations of vegetation by climate change will all be explicable by differential disturbance, differential availability of species, and differential assorting of species as they respond to one another and the physical environment. For example, seasonal changes in communities may be the result of life cycle differentiation through the year of a suite of species already present, while many successions will reflect subsequent waves of

immigration. Thus, although the specific mix of causes operating at these different scales can vary, the same conceptual framework can be employed for cause at any scale. This is one reason we have used a neutral definition of succession.

These two propositions identify the conceptually complete general causes of succession, and suggest that the same suite of causes operate for intra-annual to paleoecological scales. One insight to emerge from the three-cause framework is that models of community or ecosystem structure may be partitioned along these three axes. Some models may emphasize disturbance, while others may emphasize the sorting processes among organisms such as competition. As mentioned earlier, Eliot's (2007) analysis of the Clements/Gleason divide shows just such a difference in emphasis. Ecology as a whole seems to employ a conceptual phase space defined by axes of disruption versus sorting. This is exemplified by the controversy (e.g., Gravel et al. 2006) over Hubbell's (2001) neutral models of succession in which temporal patterns are driven not by limiting similarity but by the vagaries of arrival at a site. Sorting, on the other hand, reflects differential species availability and differential species performance. Because succession theory incorporates both sorting and disruption, it is hospitable to models of assembly, lottery dispersal, response to disturbance intensity and frequency, and neutrality.

PROPOSITIONS ABOUT COMMUNITY TRAJECTORIES

It is important to separate process and outcome based ideas of succession, as one necessarily generates the other. Furthermore, process approaches are motivated by explaining how a phenomenon occurs, while outcome based approaches focus on end points. Process approaches may be especially applicable to nonequilibrium systems, while outcome based approaches may be appropriate to equilibrium situations. In addition, outcome based approaches must account for the contingencies of particular times and places. This section presents four propositions about the outcomes of the causes and mechanisms of succession. These propositions suggest the form and general content that specific models of succession must account for to capture the richness of successional outcomes.

Proposition 8: Net successional outcomes of interaction between individuals are A) tolerance or no effective interaction, B) inhibition, or C) facilitation of compositional or structural change.

239

Succession was initially *defined* as a progressive and gradual change based on the facilitation of later species by earlier dominants. This expectation was not universally supported by succession studies, especially those encompassing the fine scale and long-term (Niering 1987; Pickett 1989). As a result, the roster of successional turnovers was extended to become logically exhaustive by including tolerance and inhibition as potential outcomes (Connell & Slatyer 1977). Although inhibitive interactions had been recognized for a long time, they were initially neglected in the classical successional theory (Egler 1954). Contemporary animal ecology incorporates such dynamics via the concept of priority effects (Drake et al. 1996). Likewise, the role of physiological tolerance or life cycle characteristics was recognized in some studies, but these neutral factors were not included as part of the conceptual apparatus of successional theory (Keever 1950). Incorporating the multiple causes of Connell and Slatyer (1977) into successional theory was a crucial advance. Accepting neutrality was delayed by the difficulty of experimentally distinguishing it from niche based interaction.

This advance required one additional insight to emerge fully as a tenet of successional theory. Facilitation, inhibition, or neutrality of species turnover are in fact net effects of specific interactions and mechanisms (Pickett et al. 1987a, b). A given net effect can reflect a mixture of specific mechanisms, so it is important to discern the causes that operate in specific successional turnovers. Moreover, a given successional sequence will probably include all three net effects of facilitation, inhibition, or neutrality and tolerance (Hils and Vankat 1982; Armesto and Pickett 1986). It is incorrect to expect the successional dynamics of a system to operate via only one kind of net effect, though one may dominate. Different net effects may occur at different times, or at different locations within a dynamic community. However, they may well occur simultaneously. For example, while some pioneering annual plants may facilitate the invasion and survival of other species, biennial dominance in the second year of succession in abandoned fields is simply a reflection of their life cycles (Keever 1983).

Proposition 7: The species composition of a site tends to equilibrate with the prevailing environment of that site.

One of the fundamental principles of ecology is the tendency of biotic composition and structure to reflect the prevailing environment (Chapin et al. 2002). This idea is exemplified by the characteristic faunal and floral composition of biomes as a reflection of climate (Clements

and Shelford 1939), or of Holdridge's (1947) life zones as a response to temperature, rainfall, and altitude, or of C. Hart Merriam's (Merriam and Steineger 1890) attitudinally delimited zones on mountains. This generalization emerged from biogeography, and was well accepted by the time ecology took root as a distinct discipline.

The idea was translated into ecology as the idealized concept of climatic climax and its local variants. The justification for this concept was seen as problematical, and after long debate, was rejected by ecologists (Botkin and Sobel 1975; Johnson 1979). However, the coarse scale tendency of species lists and vegetation structures to sort at continental and regional scales is clear (Clark and McLachlan 2003). As an ideal, it is thus still possible to posit that vegetation tends to equilibrate with the prevailing environment. At its most demanding, this proposition takes the basic assumption about plant-environment relationships in plant ecology, and applies it to a temporal context. At its most liberal, the proposition suggests that the contingent pathways observed in real, messy successions will still reflect the environmental limits that prevail over the course of time in a given location. It can be considered one of the boundary laws of community dynamics: No community will exist outside the environmental limits of its component species and ecological interactions. Thus, Proposition 7 on one level reflects a fundamental assumption in the paradigm of organism-environment relations, while on another level it motivates experimental or model tests to evaluate the impact of factors and events that can deflect the composition or architecture of a community from the ideal expected in a given environment. Relic communities are a case in point that seems to violate the proposition, but they do so only because of the lag in compositional change resulting from their dominance by long-lived organisms established under a different prevailing environment. The next proposition examines the possibility of narrowing the ideal still further, excluding disturbance, heterogeneity, or other factors that can in reality determine vegetation composition and structure. We move to a proposition that focuses on the sorting component of vegetation dynamics.

Proposition 8: The specific form of a successional trajectory is contingent upon starting conditions, and the stochasticity of invasion and controls on species interactions. The contingencies that can operate in succession are many. Some of these are set by the conditions that exist at the start of the succession, say a wet versus a dry year, or the species that happen to be spared by a specific disturbance event. Others are the result of events and

conditions that may appear stochastically at various times through succession, such as fine scale physical disturbance events or outbreaks of herbivores. Contingencies may exist in A) the extent of environmental contrasts or resource gradients that can exist at a site through time, B) the adaptive repertoire, including plasticity, contained within the species seeds, eggs, or clonal fragments present; C) the type, size, and timing of disturbances, D) climatic shifts in the resource base, E) the order in which species colonize, and F) the external landscape context.

This proposition accepts that succession is contingent on a large suite of interacting drivers, generating the inherent variability noted for successional systems. Any tendency for the composition or structure of an assemblage to equilibrate with an environment at a particular scale (Proposition 7) is constrained by the strength of disequilibrating forces occurring at that scale (Huston 1994). Thus, the utility of models of an environmentally-determined, equilibrated species assemblage is a particular research choice, and not a universal feature of all successions. Certain models may assume an edaphic successional "attractor" or an "extremum" condition. Whether the ideal environmental equilibrial community can be used as an end point of sorting depends on whether the disrupting forces are sufficiently weak relative to the prevailing environmental filters over the time period examined. Environments in which the conditions change during the time environmental sorting takes place will probably not generate a stable attractor or endpoint. If the disturbance regime, adjacent landscape structure, species pool, or biophysical template change, successional trajectories may also differ from one period to the next. Such variety can, of course, lead to the potential for multiple stable states (Holling 1992).

Some models recognize that plant assemblages may respond more to the patterns of disturbance and species availability than to uninterrupted species interaction. In such non-equilibrium situations, succession will be dependent upon landscape context, and the variation between sites that may reflect different successional states or histories. This means that succession cannot be understood by focusing only on a single site or set of local conditions (DeGraaf and Miller 1997; Ejrnaes et al. 2003; Pickett 1976). The dynamics and structure of adjacent sites may influence a given succession by serving as a source of potential colonists or new species with low resource tolerances.

One contingency that may determine successional trajectories is the length of the resource gradient that can exist over time. Different habitats can show long or short temporal gradients of compositional change based on the resources levels and stresses that dominate the

region. The length of the resource gradient acquires further biological significance based on the maximum level of resources that can be accumulated. Species assembly in low resource sites may exhibit a continued species accumulation with time as resources accumulate, while in high resource sites, species richness may peak at intermediate periods of succession (Auclair and Goff 1971).

The order of species colonization is important to the details of successional trajectories because it may determine the order of facilitative, inhibitory, or neutral interactions. Priority effects are especially obvious in successions on sites having low resources (Lockwood 1997).

Proposition 9: Succession produces temporal gradients of the physical environment, biotic communities, and the interaction of the two.

Resources and regulators change with succession. Succession is a feedback between a species pool and a complex biotic-abiotic environmental gradient (Pastor and Post 1986). Although the physical environment sets the initial resource base, physiological stress, and regulator conditions, these are subsequently modified by the species and their interactions in the succession. Therefore, succession generates a complex environmental gradient through time.

A long term ecosystem successional gradient is illustrated by the trajectory of nutrient pools through time (Vitousek 2004). For nutrients, such as P, that are borne in the substrate in which soil development occurs, the pattern over the very long term is one of leaching and loss (Figure 9.3). For nutrients, such as C and N, whose concentration in the ecosystem depends on biological metabolism, the amount will at first increase, and then decrease in the ecosystem. This means that over the very long term, nutrient limitation in ecosystems will shift from N to P. When subsets of these time frames are examined following discrete disturbances, shorter successional patterns can be discerned. The nutrient dynamics of these shorter successions will be determined by the relative degree of N and P limitation that has been generated by the longer succession and are in effect for that time span. For example, long-term leaching of nutrients following glacier retreat at a regional scale may set fertility levels, while succession in individual abandoned agricultural fields in the area may generate local dynamics.

This proposition suggests that while for many ecological purposes considering resource gradients to be independent and fixed is appropriate, for succession studies, these gradients should be considered a dynamic part of the system. The outcomes of species interaction can create environmental legacies affecting subsequent dynamics. The origin of the ecosystem concept was stimulated by the insight that succession involves the reciprocal interaction of the physical and the biological complexes at a site, following the logic of this proposition, rather than some ill defined organismal tendencies (Tansley 1935).

STRUCTURE AND OUTPUT OF SUCCESSION THEORY

To this point, we have examined the individual components, or propositions, of succession theory. However, a theory is in fact an integrated conceptual and empirical structure (Pickett et al. 2007). We will relate the various propositions and the roles of the theory components they represent in this section.

We have already used one organizational component as a framework for successional theory – the general causal hierarchy (Figure 9.1). This consisted of a statement of the domain, an identification of the general causes of succession, and subsets of more specific causes, that is, the constraints and mechanisms that drive succession in particular models or situations. The specific models, which in the minds of many are unfortunately the entirety of theory, are tools to put those causes into action. Models select the mechanisms or interactions that will be related in particular studies, set the spatial and temporal boundaries of interest, specify the inputs and results that will be addressed, and limit the kinds of outcomes. Laws are statements that specify the limits on interactions, the nature of trajectories, or indicate boundary conditions that cannot be violated. Models and laws are closely related, as they are both intended to describe aspects of the behavior of the system of interest (Pickett et al. 2007). However, models are often used to indicate how the general laws apply to specific real or hypothetical situations. Furthermore, models tend to be more complex than laws, as they address many dimensions or features of a system, while laws often focus on a few simplified or general factors. We discuss successional laws and models as organizing and operationalizing tools of succession theory.

Successional laws

When present, an important organizational tool for general theory is its suite of laws. The law of succession parallels the laws of evolution in structure and use (Pickett and Cadenasso 2005; Pickett et al. 1994). The law of succession (Pickett and McDonnell 1989) has a universal conditional form, which indicates what the possible causes of succession are. The law of

succession states that 1) if sites are differentially available; and 2) species are differentially available at those sites; or 3) species perform differentially at those sites, then the aggregate community structure or composition will change through time. A "zero force" form is possible by casting the statements in terms of the absence of the phenomenon each identifies as resulting in no net change in the community. Differential availability must be conceived of to include the preexisting conditions at various sites (Figure 9.2), which will interact with the various intensities, sizes, and frequencies of disturbance to condition sites for successional change. It is a complex of conditions that constitute site availability, rather than a simple concept of a vacant site (Walker 1999). Differential species availability can include persistence through the initiating disturbance or migration from elsewhere. Hence, a patch dynamic or metacommunity perspective (Hastings, this volume; Leibold, this volume) is implied by the core processes of succession (Pickett 1976; Pickett and Rogers 1997). Finally, it is important to recognize that differential performance can be achieved based on evolved species traits, interactions at the producer trophic level, or by interaction with consumers in the form of diseases, predators, and herbivores. Differential performance therefore incorporates the potential for frequency dependent, density dependent, and stochastic constraints on limiting similarity, as well as limiting similarity itself.

Predictions, Expectations, and Hypotheses

The propositions outlined above are primarily in the realm of principles that lay out the high level, general structure of community change theory (Jax 2006). They focus on process, and the modes by which successional outcomes occur. They provide the raw materials for the specific models from which expectations about succession will emerge. Specific models will usually address some narrower domain than the entire theory of succession, or will focus on a subset of the kinds of interaction among the entire roster possible (Figure 9.1). In other words, the specific models of succession are nested within the causal repertoire and address a certain subset of successional relationships. They operationalize the causal repertoire. A general template for such models is illustrated in Figure 9.4.

The fact that specific models of succession will usually address subsets of the causal and contextual universe of succession means that classification of cases is required. We have already pointed to one kind of classification – emphasis on disruption of species interactions as the

dominant driver of community composition, versus emphasis on sorting among species as the dominant driver. This classification isolates the process of differential site availability from the remaining two differentials of successional process – species availability and species performance. One could further remove differential species availability by constructing models assuming equal access by all species or conducting experiments in which species were made uniformly available. Experimental studies have recently begun to examine the role of interactions between vertebrate consumers and dispersers with landscape structure as controls on differential species availability (e.g., Ostfeld, et al. 1999; Cadenasso and Pickett 2000; Meiners and LoGiudice 2003).

One of the most familiar classifications of succession is the dichotomy between primary and secondary successions (del Moral 1993; Walker 1999). The main difference between these two classes is in the resource availability of the sites representing each type. Primary successions are defined as those beginning on a new substrate that is often low in biologically mediated resources such as C and N, but perhaps relatively high in pools of substrate-based nutrients, such as P (Walker 1993). Primary successions can differ in the forms and identities of the nutrients available in the substrate based on the origin of the rocks or sediment. For example certain volcanic parent materials can yield fertile soils. Primary successions also are expected to largely lack resident propagules or survivors compared to secondary successions, in which clonal fragments, dormant seeds, or animal resting stages commonly survive the initiating disturbance (Myster 2008). This classification identifies resource pools, biological resource legacies, and biological propagule legacies as three key initial conditions for succession models. It is mechanistically safer to identify the specific resource levels and species availabilities in different succession than to rely on the primary and secondary dichotomy to adequately characterize them (Figure 9.5). Some successions expected to behave as primary in fact possessed some attributes, such as survival of propagules or resource legacies through the disturbance (del Moral 1993; Walker and del Moral 2003).

Succession models and models that are relevant to succession take on a vast array of forms. Lottery models are in essence models emphasizing the role of site disruption and differential availability in contrast to emphasis on interactive species sorting (Chesson 1991). Most successional modeling focuses on the process of sorting, and therefore emphasize differential species performance. For example, Markov matrices assume stationarity in the transition probabilities from earlier to later communities (Usher 1979, 1992). Gap based models are driven by the differential shade tolerance of species that participate in the succession (Horn et al. 1989). Individual based models have recently become the predominant strategy for modeling succession, and some now incorporate neighborhood effects of dispersal or neighborhood effects on soil and above ground resources (Canham et al. 2003).

Modeling succession hinges on whether the process or the end point is emphasized. Compositional or biogeochemical equilibria after long periods of sorting or resource partitioning between production, maintenance, and storage can be useful points of reference (Odum 1969). These states are potential end points of succession. In contrast are successional models that focus on situations that can be interrupted by disturbance before idealized sorting can be completed (Peet 1992), or situations in which not all species are present equally (Hastings 1980; Petraitis et al. 1989; Whittaker and Bush 1993), or in which consumers or stress factors intervene strongly to re-sort species (Olff et al. 1999).

There are a number of idealized expectations that apply to succession, *ceteris paribus*. One is the intermediate disturbance hypothesis, in which high intensities or frequencies of disturbance are assumed to permit the persistence of only a limited pool of species that are well adapted to the effects of disturbance (Hastings 1980; Petraitis et al. 1989). As disturbance frequency or intensity is relaxed, the number of species that can coexist in the site increases due to the release of the disturbance constraint. Ultimately, however, competition and other biotic interactions begin to impose a limit based on the exclusion of species by superior competitors. This relationship assumes a tradeoff between competitive ability and tolerance of disturbance or stress. The theoretical scheme of Grime (1979; Grime and Hodgson 1987) makes the role of stress explicit, and suggests that community composition reflects the species tolerance of disturbance of disturbance of stress on another, and differentiation between complexity and reality to the expectations of the intermediate disturbance hypothesis. The intermediate disturbance hypothesis has been difficult to discriminate in the field due to the many factors that can function simultaneously with physical and biological tolerances.

Another *ceteris paribus* expectation of succession is the intermediate richness hypothesis. The expectation here is that during intermediate phases of succession in productive sites, species richness will be highest (Loucks 1970). Like the intermediate disturbance hypothesis, the intermediate richness hypothesis assumes that relatively fewer species are adapted to the stresses of early successional habitats than are adapted to the more moderate physical conditions that are generated by the ecological engineering as species invade, grow, and coalesce their biotic effects. However, under the physically more moderate conditions that often begin to appear as succession proceeds, more biotically competitive species dominate, and exclude the pioneers and earlier dominants. This process will result in lowered diversity in later succession.

The intermediate diversity or richness hypothesis led the prominent plant population biologist John Harper (1967) to ask why species didn't evolutionarily climb its own successional tree? That is, why didn't species adapt to the changing conditions during succession? There are two reasons for this provocative expectation not borne out. First, all species are to some extent fugitive (Hanski 1995). That is, they migrate and abandon sites where they are less well adapted and establish in sites where they are better adapted. Plants probe their potential environments via dispersal, and mobile animals actively probe sites. Second, the contrasts between adaptation to high resource levels versus adaptation to low resource environments involves both physiological and architectural trade offs that are difficult to shift evolutionarily over the short term (Bazzaz 1979, 1983). Individual genera may have species that are arrayed along an adaptive successional gradient, but rarely does an individual species contain within it the genetic scope – including plastic capacity – to address such a broad array of conditions (Tilman 1991). Putting these two generalizations together explains why species exploit shifting successional mosaics in landscapes rather than genetically adjust to the sere at any one point in space (Pickett 1976).

A modification of the intermediate richness hypothesis recognizes the constraints of xeric systems. Rather than an ultimate decrease in diversity over the span of several centuries as seen in mesic systems, xeric plant communities would be expected to show continuing increase in diversity (Auclair and Goff 1971). Such secondary successional systems may in fact be responding more like primary succession due to the stress imposed by low levels of resources available. Over very long time periods, however, such systems would be expected to show the diversity decline, just as the very oldest primary successions do (Vitousek 2004).

Ecosystem parameters are also expected to change during succession. Like the ideal diversity expectations, these are based on the evolutionary tradeoffs that species exhibit. The relevant tradeoff here is between the high metabolism associated with using freely available and uncontested resources and the slow metabolism and high storage capacity associated with

exploiting contested and low levels of available resources (Odum 1969). Metabolic rate is also inversely associated with length of life. Putting these assumptions together suggests that early successional ecosystems will exhibit high productivity, low storage in soil pools, and a relatively low proportion of its total metabolism in respiratory activity. Hence, ecosystem metabolism is expected to have greater production than respiration early in succession, and to eventually come to a point after which production and respiration balance (Bormann and Likens 1979). Food webs should show concomitant shifts from grazer to decomposer dominance over this span (Quetier et al. 2007). An additional expectation is that system sensitivity to external disturbance would increase as the species invest more in a complex structure that has low productive potential per unit time (Gunderson 2000).

The current controversy and excitement concerning neutral theory (Gravel et al. 2006; Hubbell 2001) relates to succession theory. The neutral model posits that species have identical properties, and that local coexistence can be predicted by localized, as opposed to global, dispersal dynamics (Bell 2001). Mechanisms of coexistence can be divided among those that 1) are based on niche partitioning and competitive tradeoffs, 2) frequency dependence driven by prey-specific consumers, 3) limited recruitment to local habitats due to restricted dispersal, or 4) a dynamic equilibrium of speciation and extinction, i.e., neutral drift (Chave 2004). The two broad kinds of models based on neutrality or differentiation cannot be distinguished based on pattern data from the field (Chase 2005; Gravel et al. 2006). Indeed, Pueyo et al. (2007) showed that there a large number of plausible models between neutrality and idiosyncratic species behavior that can generate the same diversity outcomes. An emerging view is that local community organization is open to landscape or metacommunity influences, and that niche partitioning and neutrality can be usefully treated as ends of a continuum (Gravel et al. 2006). Which kind of mechanism -- competitive exclusion based on niche difference, or stochastic exclusion – operates in a particular situation depends upon whether immigration can act to prevent limiting similarity (Gravel et al. 2006). This logic parallels the division of successional causes between those that disrupt community composition and open up sites, versus those that operate through sorting species along the complex ecosystem gradient of succession. The theoretical option of community control by migration was opened in the 1950s (Skellam 1951), and has matured through the intellectually parallel streams of island biogeography, metapopulation theory, patch dynamics, and metacommunity theory (Bell 2001; Leibold et al.

2004; Pickett and Rogers 1997). Contemporary succession offers a framework to explicitly link these seemingly disparate views (Figure 9.6).

STRUCTURE OF SUCCESSION THEORY: A SUMMARY AND LINKAGES

Succession or community dynamics theory is rich and well connected with other biological and ecological theories. One of our tasks is to relate community dynamics theory to the fundamental principles of ecological theory (Scheiner and Willig Chapter 1).

1. Heterogeneous distributions of organisms. Succession is an case of heterogeneous distributions of organisms and their aggregations in both time and space. Through time, succession illustrates either architectural or compositional changes in communities. The temporal distribution of dominance is heterogeneous, so that individual species abundances show overlapping or idiosyncratic patterns at the fine scale. Often long lags precede or follow peak abundances, having apparently stochastic minor fluctuations. Succession itself becomes a higher order source of heterogeneity, as successions that start at different times or occupy different sites across landscapes and regions. The existence across landscapes of communities, patches, and ecosystems that have different successional origins, ages, and characteristics is a dimension of environmental heterogeneity that affects other ecological patterns and processes.

2. Environmental interactions of organisms. The heterogeneous distributions of organisms during succession are the result, in part, of the interactions between the organisms and the environment that prevails at different times. Succession can be conceived as a temporal gradient of environmental change, driven by the different organisms that predominate through time. The feedback between organisms and environment is a key aspect of successional pattern and process.

3. Contingency: History and Stochasticity. This fundamental principle of ecology is compellingly applied to succession. The initial and boundary conditions of a particular community determine much about its subsequent temporal trajectory. Stochasticity appears as some organisms happen to survive a catastrophic disturbance, or as the organisms that happen to be within close dispersal range dominate the disturbed site, or when stresses, resources, and disease agents infiltrate unpredictably from nearby landscape elements. Contingency is well illustrated by the increasing role of accidentally or intentionally introduced plant, animal, and microbial species in the trajectories of various successions. Likewise, historical contingencies

appear in those current communities where composition and changes are conditioned by the past management by humans. These include such legacies as the fire or hunting practices of indigenous populations no longer dominant, or the nature of more recent, but still past agricultural practices on succession. The order of establishment is an important contingency in specific communities. In addition, contingencies also take the form of stochastic interactions with climate cycles and disturbance events.

4. Heterogeneity of environmental conditions. Succession responds to differences in environmental conditions from place to place. The classical distinction between primary and secondary successions is an expression of heterogeneous resource levels and stocks of resting stages, seeds, or clonal fragments in disturbed sites. Even within secondary successions, specific trajectories differ along spatial environmental gradients as a result of differences in soil nutrients, moisture, and interacting consumer populations.

5. Finite and heterogeneous resources. The limitation of resources is key to understanding succession. Individual limiting resources change in abundance through succession and ratios of above ground to below ground limits change. Not only is the external environment a source of limitation and variation, but the internal limitation of resources within each organism is an important explanatory principle in succession. Internal limits require that resources allocated to one structure or process are not available of other structures and processes. This principle of allocation is the basis for understanding why genotypes – whether within or between species – are usually not capable of exploiting widely contrasting successional environments. Differences in capacities for dispersal, competition, mutualism, defense and food or resource acquisition all depend on this fundamental limitation of available and assimilated resources by organisms.

6. Birth and death. Succession occurs because organisms are mortal. Regardless of their inherent patterns of senescence, they may be killed by predation, disease, or physical disturbances. When large numbers of organisms occupying a site are killed, the mortal event is almost always followed by colonization, interaction, and sorting of other individuals or species. Even when a single canopy tree dies in a forest, either brought low by disturbance or by senescence, succession can be affected. In mesic forests, more shade tolerant trees can ascend to the canopy after mortality of prior canopy dominants that had reached the canopy under less light limiting conditions. If there is no clear heir apparent in the understory, new establishment or

thinning among of a number of saplings can occur. Vegetation succession has been summarized by some as a plant-by-plant replacement process. Turnover in animal communities can also involve birth and death, but migration is a common alternative for mobile organisms. Mortality agents important to succession include both physical events and the depredations of herbivores, diseases, and predators.

7. Evolution. Evolution is the biological mechanism underlying the contrasts among species that are worked out in succession. Allocation strategies, mentioned above, are evolutionary products. Contrasts among the crown architectures of plants adapted to shady versus sunny conditions are shaped by evolution. Chemical and mechanical defenses reflect evolutionary histories with their metabolic costs and reproductive benefits. Indeed, the spatial and temporal patchwork of habitats characterized by contrasting successional composition and environment must be one of the important stages for the enactment of the evolutionary play and the spatial or interactive assortment of its products.

8. Heterogeneity of entities. In the overview of fundamental principles, this one was considered a result of the other seven (Scheiner and Willig Chapter 1). So it is with succession. Successions themselves can contribute to environmental heterogeneity in time and in space. During succession heterogeneity often accumulates, at least in certain ranges of the process, as clonal organisms, neighborhood effects, and architectural complexity of dominants establish and increase. The various contingencies experienced by specific successions can lead to great heterogeneity over space in the communities, ecosystems, and habitats that are available for organisms and in which biogeochemical processes differ.

The relevance of succession to the fundamental principles of ecology is one indication of the synthetic power of the concept. Additional aspects of its synthetic significance can be understood based on its structure.

First, ecological concepts have three dimensions – meaning, model, and metaphor (Pickett and Cadenasso 2002). We believe that neutral meanings or definitions are most generalizable, but that they require models to translate them to different cases, or to expose the interactions of the causes of the process (Jax 2006, 2007). Metaphor, the third dimension of an ecological concept, has two uses in ecological theories. Metaphors can serve as an initial stimulus for a theory, or as an image to translate the theory or its components to non-specialists or to the public. In the case of succession, the term itself is clearly a metaphor originally derived
from society. When the theory was introduced, one community was envisaged to be replaced by another that is in the aggregate better adjusted to the environment created by the previous community. However, as we have seen from the causal richness and the complex reality embodied in the propositions of contemporary succession theory, this simple image is incomplete (Table 9.1). Unfortunately, there is no single image that has emerged to take the place of the royal analogy of orderly transitions. Cooper (1926) suggested that community dynamics was better imagined as a braided stream than a swift, straight channel, which implies multiple pathways and continuous change, but at faster or slower rates depending on the nature of the particular channel. Backwaters and eddies are of course part of Cooper's image.

The contemporary theory of succession is, as noted above, causally rich. Even though Clements originally proposed a remarkably broad conception of the causes of plant community change, he left out important mechanisms such as herbivory, and seemed to set disturbance outside the process of sorting that he was most interested in (Pickett et al. 2008). Vegetation dynamics, and indeed the dynamics of all sorts of communities and at all spatial and temporal scales is driven by the differentials between available sites, the differential availability of species to a site, and their differential performance within a site. Succession is seen to take place in a dynamic spatial context, which can be identified as a landscape approach, or patch dynamics, or metacommunity dynamics. Succession therefore has a rich, spatially explicit repertoire of causes.

Several model templates exist in succession. The one most closely related to the causal hierarchy shows how the subsets of those general processes might act to filter the species interacting through time at a site (Figure 9.4). Other model templates exist, depending upon whether the aggregate assemblage is the model target, as it is in Markov models (Horn 1975), or gap replacement is the focus (Shugart 1984), or individual woody plants in their spatial neighborhoods are the target, as in the Sortie model (Pacala et al. 1993). The particular models represent the fruits of various of the twigs of successional causes and

mechanisms – to turn the causal hierarchy (Figure 9.1) on its side, and place the three general causes closest to the root of the process. The models combine the specific mechanisms to precisely determine how individual organisms interact with one another and with the physical environment to generate community change through time. The large computing power now available permits spatially explicit, individual based models to address complex mixes of

dispersal, resource levels, competition, and survival, for example (e.g. Clark 2007). No longer is it necessary to decide which single factor drives succession based on mutually exclusive, alternative hypotheses. Rather, discovering the mix of factors and interactions that come into play has great power for understanding community trajectories, and for examining how changing land use and climate (Bazzaz 1986; Bazzaz and Sipe 1987) can modify trajectories. The plethora of models is a response to rejecting the unrealistic assumption of stationary contexts and conditions for most successions.

Succession has remarkable powers to synthesize different ecological perspectives and applications (Figure 9.7). Oddly, this seems only rarely to be recognized, perhaps because the terminology used by the otherwise relevant disciplines and community dynamics theory overlap so little. For example, the disparate concerns of invasion ecology, focusing on conditions in potentially invasible sites, or characteristics of the potential invading species, seems locked in arguments of alternative hypotheses when in fact, some complex mixture of mechanisms can operate, as suggested by the successional framework (Davis et al. 2005). Succession has the power to unify aspects of many other ecological perspectives as well. For example, the spatially explicit concerns of landscape ecology are central to the workings of disturbance and the distribution of propagules between disturbed and successional sites. Assembly rules are expressions of the interactions within communities which lead to successional sorting. Indeed, community assembly is preeminently a temporal process (Keddy 1992). Assembly emerges from the subtleties of autecology and physiological ecology, which expose adaptations that clearly contribute to the local coexistence of organisms (Bazzaz 1986). Organismal physiology, body plans, architecture and morphology, and the timings of life cycle events that are summed up in the unfortunately neglected term, autecology, are the stuff of sorting along successional gradients. Indeed, gradient theory is one of ecology's powerful and pervasive ideas that is expressed in a successional context through time as much as in a spatial context (Austin 2005). The gradients to which species respond during succession are not independent variables. Instead, they are emergent properties resulting from the interaction of organisms with one another, and with the physical environment (Vitousek 2004). The ecosystem concept grew out of this recognition, and it is helpful to think of successions as complex gradients of ecosystem change. There is little of contemporary ecology that does not contribute to successional explanation, or benefit from understanding the spatial and temporal dynamics of ecological

systems. That spatial and temporal dynamic is the essence of succession. Thus, one of ecology's oldest theories turns out to continue to have broad and adaptive relevance as a synthetic tool.

Acknowledgements. We are grateful to the editors and two anonymous reviewers for their helpful comments on an earlier draft of the paper. We also acknowledge NSF Long-Term Research in Environmental Biology grant DEB-0424605.

LITERATURE CITED

- Ahl V. and Allen T.F.H. 1996. Hierarchy theory: a vision, vocabulary, and epistemology. Columbia University Press, New York.
- Armesto J.J. and Pickett S.T.A. 1986. Removal experiments to test mechanisms of plant succession in oldfields. Vegetatio 66: 85-93.
- Auclair A.N.D. and Goff F.G. 1971. Diversity relations of upland forests in the western Great Lakes area. American Naturalist 105: 499-528.
- Austin M.P. 2005. Vegetation and environment: discontinuities and continuities. In: van der Maarel E. (ed), Vegetation ecology. Blackwell Science, Malden, MA, pp. 52-84.
- Bazzaz F.A. 1979. The physiological ecology of plant succession. Annual Review of Ecology and Systematics 10: 351-371.
- Bazzaz F.A. 1983. Characteristics of populations in relation to disturbance in natural and manmodified ecosystems. In: Mooney H. A. and Godron M. (eds), Disturbance and ecosystems: components of responses. Springer-Verlag, New York, pp. 259-275.
- Bazzaz F.A. 1986. Life history of colonizing plants: some demographic, genetic, and physiological features. In: Mooney M. A. and Drake J. A. (eds), Ecology of biological invasions of North America and Hawaii. Springer-Verlag, New York, pp. 96-110.
- Bazzaz F.A. and Sipe T.W. 1987. Physiological ecology, disturbance, and ecosystem recovery. In: Schulze E. D. (ed), Potentials and limitations of ecosystems analysis. Springer-Verlag, New York, pp. 203-227.
- Begon M., Harper J.L. and Townsend C.R. 1996. Ecology: individuals, populations, and communities. Blackwells, Oxford.
- Bell G. 2001. Neutral macroecology. Science 293: 2413-2418.
- Bever J.D. 2003. Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. New Phytologist 157: 465-473.
- Biox D., Quintana X.D. and Moreno-Amich R. 2004. Succession of the animal community in a Mediterranean temporary pond. Journal of the North American Benthological Society 23: 29-49.
- Bormann F.H. and Likens G.E. 1979. Patterns and processes in a forested ecosystem. Wiley and Sons, New York.

- Botkin D.B. and Sobel M.J. 1975. Stability in time-varying ecosystems. American Naturalist 109: 625-646.
- Boucher D., Jardillier L. and Debroas D. 2005. Succession of bacterial community composition over two consecutive years in two aquatic systems: a natural lake and a lake-reservoir. FEMS Microbial Ecology 55: 79-97.
- Bowers M.A. 1997. Influence of deer and other factors on an old-field plant community. In: McShea W.J., Underwood H.B. and Rappole J.H. (eds), The science of overabundance. Deer ecology and population management. Smithsonian Institution Press, Washington, D.C., pp. 310-326.
- Brand T. and Parker V.T. 1995. Scale and general laws of vegetation dynamics. Oikos 73: 375-380.
- Cadenasso M.L. and Pickett S.T.A. 2000. Linking forest edge structure to edge function: mediation of herbivore damage. Journal of Ecology 88: 31-44.
- Cadenasso M.L., Pickett S.T.A., Weathers K.C., Bell S.S., Benning T.L., Carreiro M.M. and Dawson T.E. 2003. An interdisciplinary and synthetic approach to ecological boundaries. BioScience 53: 717-722.
- Canham C.D., Cole J.J. and Lauenroth W.K. (eds). 2003. Models in ecosystem science. Princeton University Press, Princeton.
- Chapin F.S., III, Matson P.A. and Mooney H.A. 2002. Principles of terrestrial ecosystem ecology. Springer-Verlag, New York.
- Chase J.M. 2005. Towards a really unified theory for metacommunities. Functional Ecology 19: 182-186.
- Chave J. 2004. Neutral theory and community ecology. Ecology Letters 7: 241-253.
- Chesson P.L. 1991. Stochastic population models. In: Kolasa J. and Pickett S. T. A. (eds), Ecological heterogeneity. Springer-Verlag, New York, pp. 123-143.
- Clark J.S. 2007. Models for ecological data: an introduction. Princeton University Press, Princeton.
- Clark J.S. and McLachlan J.S. 2003. Stability of forest biodiversity. Nature 423: 635-638.
- Clements F.E. 1916. Plant succession: an analysis of the development of vegetation. Carnegie Institution of Washington, Washington.
- Clements F.E. and Shelford V.E. 1939. Bio-ecology. Wiley, New York.

- Connell J.H., Noble I.R. and Slatyer R.O. 1987. On the mechanisms producing successional change. Oikos 50: 136-137.
- Connell J.H. and Slatyer R.O. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. American Naturalist 111: 1119-1144.
- Cooper W.S. 1926. The fundamentals of vegetation change. Ecology 7: 391-413.
- Davidson D.W. 1993. The effects of herbivory and granivory on terrestrial plant succession. Oikos 68: 23-35.
- Davis M.A., Pergl J., Truscott A.-M., Kollmann J., Bakker J.P., Domenech R., Prach K., Prieur-Richard A.-H., Veeneklaas R.M., Pysek P., del Moral R., Hobbs R.J., Collins S.L., Pickett S.T.A. and Reich P.B. 2005. Vegetation change: a reunifying concept in plant ecology. Perspectives in Plant Ecology, Evolution and Systematics 7: 69-76.
- Dayton P.K. 1975. Experimental evaluation of ecological dominance in a rocky intertidal algal community. Ecological Monographs 45:137-159.
- DeGraaf R.M. and Miller R.I. 1997. The importance of disturbance and land-use change in New England: implications for forested landscapes and wildlife conservation. In: DeGraaf R. M. and Miller R. I. (eds), Conservation of faunal diversity in forested landscapes. Chapman and Hall, New York, pp. 3-35.
- del Moral R. 1993. Mechanisms of primary succession on volcanoes: a view from Mount St. Helens. In: Miles J. and Walton D. W. H. (eds), Primary succession on land. Blackwell Scientific Publications, Boston, pp. 79-100.
- Dobson A. & Crawley M. 1994. Pathogens and the structure of plant communities. Trends in Ecology and Evolution 9: 393-398.
- Drake J. A., Huxel G.R. and Hewitt C.L. 1996. Microcosms as models for generating and testing community theory. Ecology 77: 670-677.
- Egler F.E. 1954. Vegetation science concepts. I. Initial floristic composition a factor in oldfield vegetation development. Vegetatio 4: 412-417.
- Ejrnaes R., Hansen D.N. and Aude E. 2003. Changing course of secondary succession in abandoned sandy fields. Biological Conservation 109: 343-350.
- Eliot C. 2007. Method and metaphysics in Clements's and Gleason's ecological explanations. Studies in history and philosophy of biological and biomedical sciences 38: 85–109.

- Farrell, T. M. 1991. Models and mechanisms of succession: an example from a rocky intertidal community. Ecological Monographs 61: 95–113.
- Gleason H.A. 1917. The structure and development of the plant association. Bulletin of the Torrey Botanical Club 44: 463-481.
- Glenn-Lewin D.C. 1980. The individualistic nature of plant community development. Vegetatio 43: 141-146.
- Glenn-Lewin D.C., Peet R.K. and Veblen T.T. (eds). 1992. Plant succession: theory and prediction. Chapman and Hall, New York.
- Gravel D., Canham C.D., Beaudet M. and Messier C. 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecology Letters 9: 399-409.
- Grime J.P. 1979. Plant strategies and vegetation processes. John Wiley & Sons, New York.
- Grime J.P. and Hodgson J.G. 1987. Botanical contributions to contemporary ecological theory. New Phytologist 106: 283-295.
- Gunderson L.H. 2000. Ecological resilience in theory and application. Annual Review of Ecology and Systematics 31: 425-439.
- Hanski I. 1995. Effects of landscape pattern on competitive interactions. In: Hansson L., Fahrig L. and Merriam G. (eds), Mosaic landscapes and ecological processes. Chapman and Hall, New York, pp. 203-224.
- Hansson L. and Angelstam P. 1991. Landscape ecology as a theoretical basis for nature conservation. Landscape Ecology 5: 191-201.
- Harper J.L. 1967. A Darwinian approach to plant ecology. Journal of Ecology 55: 247-270.
- Hastings A. 1980. Disturbance, coexistence, history and competition for space. Theoretical Population Biology 18: 363-373.
- Hils M.H. and Vankat J.L. 1982. Species removals from a first-year old-field plant community. Ecology 63: 705-711.
- Hobbs R.J. and Walker L.R. 2007. Old field succession: development of concepts. In: Cramer V.A. and Hobbs R. J. (eds), Old fields: dynamics and restoration of abandoned farmland. IslandPress, Washington, D.C., pp. 17-30.
- Holdridge L.R. 1947. Determination of world plant formations from simple climatic data. <u>Science</u> 105: 367-368.

- Holling C.S. 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecological Monographs 62: 447-502.
- Horn H.S. 1974. The ecology of secondary succession. Annual Review of Ecology and Systematics 5: 25-37.
- Horn H.S. 1975. Markovian properties of forest succession. In: Cody M. L. (ed), Ecology and evolution of communities. Harvard University Press, Cambridge, pp. 196-211.
- Horn H.S., Shugart H.H. and Urban D.L. 1989. Simulators as models of forest dynamics. In: Roughgarden J., May R. M. and Levin S. A. (eds), Perspectives in ecological theory. Princeton University Press, Princeton, pp. 256-267.
- Hubbell S.P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.
- Huston M.A. 1992. Individual-based forest succession models and the theory of plant competition. In: DeAngelis D. L. (ed), Populations and communities: an individual-based perspective. Chapman and Hall, New York, pp. 408-420.
- Huston M.A. 1994. Biological diversity: the coexistence of species on changing landscapes. Cambridge University Press, New York.
- Inouye R.S., Allison T.D. and Johnson N.C. 1994. Old field succession on a Minnesota sand plain: effects of deer and other factors on invasion by trees. Bulletin of the Torrey Botanical Club 121: 266-276.
- Jax K. 2006. Ecological units: definitions and application. Quarterly Review of Biology 81: 237-258.
- Jax K. 2007. Can we define ecosystems? on the confusion between definition and description of ecological concepts. Acta Biotheoretica 55: 341-355.
- Jax K., Jones C. and Pickett S.T.A. 1998. The self-identity of ecological units. Oikos 82: 253-264.
- Johnson E.A. 1979. Succession, an unfinished revolution. Ecology 60: 238-240.
- Keddy P.A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3: 157-164.
- Keever, C. 1950. Causes of succession on oldfields of the Piedmont, North Carolina. Ecological Monographs 20:229-250.

- Keever C. 1979. Mechanisms of plant succession on old fields of Lancaster County, Pennsylvania. Bulletin of the Torrey Botanical Club 106: 299-308.
- Keever C. 1983. A retrospective view of old-field succession after 35 years. American Midland Naturalist 110: 397-404.
- Keller J. K., Richmond M.E. and Smith C.R. 2003. An explanation of patterns of breeding bird species richness and density following clearcutting in northeastern USA forests. Forest Ecology and Management 174:541-564.
- Kingsland S.E. 2005. The evolution of American ecology, 1890-2000. Johns Hopkins University Press, Baltimore.
- Klironomos J. N. 2002. Feedback with soil biota contributes to plant rarity and invasiveness in communities. Nature 417: 67-70.
- Leibold M.A., Holyoak M., Mouquet N., Amarasekare P., Chase J.M., Hoopes M.F., Holt R.D., Shurin J.B., Law R., Tilman D., Loreau M. and Gonzalez A. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7: 601-613.
- Lockwood J.L. 1997. An alternative to succession: assembly rules offer guide to restoration efforts. Restoration and Management Notes 15: 45-50.
- Loreau, M. 1998. Biodiversity and ecosystem functioning: a mechanistic model. Proceedings of the National Academy of Science of the United States of America 95:5632-5636.
- Loucks O.L. 1970. Evolution of diversity, efficiency, and community stability. American Zoologist 10: 17-25.
- MacMahon J.A. 1980. Ecosystems over time: succession and other types of change. In: Waring R. (ed), Forests: fresh perspectives from ecosystem analyses. Oregon State University Press, Corvallis, pp. 27-58.
- MacMahon J.A. 1981. Successional processes: comparisons among biomes with special reference to probable role of and influences on animals. In: West D. C., Shugart H. H. and Botkin D. B. (eds), Forest succession: concepts and applications. Springer-Verlag, New York, pp. 277-304.
- McIntosh R.P. 1980. The relationship between succession and the recovery process in ecosystems. In: Cairns J. (ed), The recovery process in damaged ecosystems. Ann Arbor Science, Ann Arbor, pp. 11-62.

- McIntosh R.P. 1985. The background of ecology: concept and theory. Cambridge University Press, Cambridge.
- Meiners S.J., Handel S.N. and Pickett S.T.A. 2000. Tree seedling establishment under insect herbivory: edge effects and inter-annual variation. Plant Ecol. 151: 161-170.
- Meiners S.J. and LoGiudice K. 2003. Temporal consistency in the spatial pattern of seed predation across a forest old field edge. Plant Ecol. 168: 45-55.
- Meiners S.J., Cadenasso M.L. and Pickett S.T.A. 2003. Exotic plant invasions in successional systems: the utility of a long-term approach. Newton Square, PA.
- Merriam C. H. and Steineger L. 1890. Results of a biological survey of the San Francisco mountain region and the desert of the Little Colorado, Arizona. North American Fauna Report 3. U.S. Department of Agriculture, Division of Ornithology and Mammalia, Washington, D.C., 136 pp.
- Miles J. 1979. Vegetation dynamics. Wiley, New York.
- Morin, P. J. 1999. Community ecology. Blackwell Science, Oxford.
- Muller C.H. 1952. Plant succession in arctic heath and tundra in northern Scandinavia. Bulletin of the Torrey Botanical Club 79: 296-309.
- Myster R.W. (ed). 2008. Post-agricultural succession in the neotropics. Springer, New York, pp. 216-220.
- Myster R.W. and Pickett S.T.A. 1988. Individualistic patterns of annuals and biennials in early successional oldfields. Vegetatio 78: 53-60.
- Niering W.A. 1987. Vegetation dynamics (succession and climax) in relation to plant community management. Conservation Biology 1: 287-295.
- Odum E.P. 1969. The strategy of ecosystem development. Science 164: 262-270.
- Olff H., Vera F.W.M., Bokdam J., Bakker E.S., Gleichman J.M., de Maeyer K. and Smit R. 1999. Shifting mosaics in grazed woodlands driven by the alternation of plant facilitation and competition. Plant Biology 1: 127-137.
- Ostfeld R.S., Manson R.H. and Canham C.D. 1999. Interactions between meadow voles and white-footed mice at forest-oldfield edges: competition and net effects on tree invasion in oldfields. In: Barrett, G.W. (ed), Landscape ecology of small mammals. Springer-Verlag, New York, pp. 229-247.

- Pacala S.W., Canham C.D. and Silander J.A., Jr. 1993. Forest models defined by field measurements: the design of a northeastern forest simulator. Canadian Journal of Forest Research 23: 1980-1988.
- Parker V.T. 2004. Community of the individual: implications for the community concept. Oikos 104: 27-34.
- Pastor J. and Post W.M. 1986. Influence of climate, soil moisture, and succession on forest carbon and nitrogen cycles. Biogeochemistry 2: 3-27.
- Peet R.K. 1992. Community structure and persistence. In: Glenn-Lewin D. C., Peet R. K. and Veblen T. T. (eds), Plant succession: theory and prediction. Chapman and Hall, New York, pp. 103-151.
- Petraitis P.S., Latham R.E. and Niesenbaum R.A. 1989. The maintenance of species diversity by disturbance. Quarterly Review of Biology 64: 393-418.
- Pickett S.T.A. 1976. Succession: an evolutionary interpretation. American Naturalist 110: 107-119.
- Pickett S.T.A. 1989. Space-for-time substitution as an alternative to long-term studies. In: Likens G. E. (ed), Long-term studies in ecology: approaches and alternatives. Springer-Verlag, New York, NY, pp. 110-135.
- Pickett S.T.A. and Cadenasso M.L. 2002. Ecosystem as a multidimensional concept: meaning, model and metaphor. Ecosystems 5: 1-10.
- Pickett S.T.A. and Cadenasso M.L. 2005. Vegetation succession. In: van der Maarel E. (ed), Vegetation Ecology. Blackwell Publishing, Malden, MA, pp. 172-198.
- Pickett S.T.A., Cadenasso M.L. and Jones C.G. 2001. Generation of heterogeneity by organisms: creation, maintenance, and transformation. In: Hutchings M.L., John E.A. and Stewart A.J.A (eds), Ecological consequences of habitat heterogeneity, the annual symposium of the British Ecological Society. Blackwell, London, pp 33-52.
- Pickett S.T.A., Cadenasso M.L. and Meiners S.J. 2008. Ever since Clements: From succession to vegetation dynamics and understanding to intervention. Applied Vegetation Science 12: 9-21.
- Pickett S.T.A., Collins S.L. and Armesto J.J. 1987a. A hierarchical consideration of causes and mechanisms of succession. Vegetatio 69: 109-114.

- Pickett S.T.A., Collins S.L. and Armesto J.J. 1987b. Models, mechanisms and pathways of succession. Botanical Review 53: 335-371.
- Pickett S.T.A., Kolasa J., Armesto J.J. and Collins S.L. 1989. The ecological concept of disturbance and its expression at various hierarchical levels. Oikos 54: 129-136.
- Pickett S.T.A., Kolasa J. and Jones C.G. 2007. Ecological understanding: the nature of theory and the theory of nature, second edition. Academic Press, San Diego.
- Pickett S.T.A. and McDonnell M.J. 1989. Changing perspectives in community dynamics: a theory of successional forces. Trends in Ecology and Evolution 4: 241-245.
- Pickett S.T.A. and Rogers K.H. 1997. Patch dynamics: the transformation of landscape structure and function. In: Bissonette J. A. (ed), Wildlife and landscape ecology: effects of pattern and scale. Springer-Verlag, New York, pp. 101-127.
- Pueyo S., He F. and Zillo, T. 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecology Letters 10: 1017-1028.
- Quetier F., Thebault A. and Lavorel S. 2007. Plant traits in a state and transition framework as markers of ecosystem response to land-use change. Ecological Monographs 77: 33-52.
- Reynolds H.L., Packer A., Bever J.D. and Clay K. 2003. Grassroots ecology: plant-microbe-soil interactions as drivers of plant community structure and dynamics. Ecology 84: 2281-2291.
- Schweiger E.W., Diffendorfer J.E., Holt R.D., Pierotti R. and Gaines M.S. 2000. The interaction of habitat fragmentation, plant, and small mammal succession in an old field. Ecological Monographs 70: 383-400.
- Shugart H.H. 1984. A theory of forest dynamics: the ecological implications of forest succession models. Springer-Verlag, New York.
- Skellam J.G. 1951. Random dispersal in theoretical populations. Biometrika 38: 196-218.
- Smith V.G. 1928. Animal communities of a deciduous forest succession. Ecology 9:479-500.
- Sousa W.P. 1984. Intertidal mosaics: propagule availability, and spatially variable patterns of succession. Ecology 65: 1918-1935.
- Stevenson, R. J., C. G. Peterson, D. B. Kirschtel, C. C. King, and N. C. Tuchman. 1991. Densitydependent growth, ecological strategies, and the effects of nutrients and shading on benthic diatom succession in streams. Journal of Phycology. 27:59–69.

Tansley A.G. 1935. The use and abuse of vegetational concepts and terms. Ecology 16: 284-307.

Thompson M., Ellis R. and Wildavsky A. 1990. Cultural theory. Westview Press, New York.

- Tilman D. 1991. Constraints and tradeoffs: toward a predictive theory of competition and succession. Oikos 58: 3-15.
- Usher M.B. 1979. Markovian approaches to ecological succession. Journal of Animal Ecology 48: 413-426.
- Usher M.B. 1992. Statistical models of succession. In: Glenn-Lewin D. C., Peet R. K. and Veblen T. T. (eds), Plant succession: theory and prediction. Chapman and Hall, New York, pp. 215-248.
- Veblen T.T. 1992. Regeneration dynamics. In: Glenn-Lewin D. C., Peet R. K. and Veblen T. T. (eds), Plant succession: theory and prediction. Chapman and Hall, New York, pp. 152-187.
- Vitousek P.M. 2004. Nutrient cycling and limitation : Hawai'i as a model system. Princeton University Press, Princeton.
- Walker L.R. 1993. Nitrogen fixers and species replacements in primary succession. In: Miles J. and Walton D.W.H. (eds), Primary succession on land. Blackwell Scientific Publication, Oxford, pp. 249-272.
- Walker L.R. (ed). 1999. Ecosystems of disturbed ground. Elsevier, New York.
- Walker L.R. and del Moral R. (eds). 2003. Primary succession and ecosystem rehabilitation. Cambridge University Press, New York.
- Walker L.R., Walker J. and Hobbs R.J. (eds). 2007. Linking restoration and succession in theory and practice. Springer, New York.
- White P.S. and Jentsch A. 2001. The search for generality in studies of disturbance and ecosystem dynamics. Progress in Botany 62: 399-449.
- Whittaker R.J. and Bush M.B. 1993. Dispersal and establishment of tropical forest assemblages,
 Karakatoa, Indonesia. In: Miles J. and Walton D.W.H. (eds), Primary succession on land.
 Blackwell Scientific Publications, Boston, pp. 147-160.
- Wyckoff P.H. and Clark J.S. 2002. The relationship between growth and mortality for seven cooccurring tree species in the southern Appalachian Mountains. Journal of Ecology 90: 604-615.
- Yi, L., W. You, and Y. Song. 2006. Soil animal communities at five succession stages in the litter of evergreen broad-leaved forest in Tiantong, China. 1:142-150.



Figure 9.1. A hierarchy of successional causes, showing the most general causes on the left, and progressively more finely articulated mechanisms and constraints toward the right.



Figure 9.2. An idealized filter model of successional processes, with reference to components of the causal hierarchy in Figure 1. From Meiners, unpublished.



Figure 9.3. A model of phosphorus (P) dynamics over primary succession. After Vitousek (2004).



Redrawn after MacMahon (1980)

Figure 9.4. A model template for the interaction of processes in succession. Redrawn from MacMahon (1981).



Figure 9.5. Primary and secondary succession as differentiated along axes of site resource availability and source of propagules. From Pickett and Cadenasso (2005)



Resistance to Dispersal

Figure 9.6. Relationship of three theories commonly considered to be discrete, showing their relationship within a conceptual space defined by resistance to dispersal versus spatial scale of analysis. Based on Pickett et al. (2007).



Figure 9.7. Succession as a synthetic theory, showing suggested general relationships to other ecological theories.

Table 9.1. Contrasts between classical and contemporary theories of succession. The text parses this universe of contrast by focusing on principles of 1) domain, 2) fundamental actors, 3) environmental contrasts through succession, 4) trajectories that arise from the interactions among the actors and between actors and environment, and 5) methodology.

Successional Attribute	Classical	Contemporary
Domain	Time > 1 yr	Any time scale
Definition	Assumptions included	Neutral definition
Actors	Community action and	Individual action, community
	outcome	outcomes
	Plants and environmental	Plants; animal vectors,
	reaction	consumers, and engineers;
		feedbacks
Causes	Facilitation predominant cause	Multiple causes: Disturbance,
		species availability, species
		performance
	Focus within community	Includes landscape and
		historical context
Community trajectories	Idealized; directional	Contingent; probabilistic
	Vegetation conforms to	Vegetation conforms to
	climatic ideal	multiple, contingent factors
Theory structure	Verbal idealizations	Hierarchical causal repertoire
	Equilibrium, aspatial models	Individual based, spatial
		models
Methodology	Chronosequence	Long-term plots; simulation
		models
Theoretical context	Progressive evolution;	Contingent evolution; process
	organismal analogy with	focus
	endpoint focus	

Chapter10: The Equilibrium Theory of Island Biogeography Dov Sax and Steven D. Gaines

INTRODUCTION

The Equilibrium Theory of Island Biogeography (ETIB) is arguably the single most influential theory in the study of geographic patterns of diversity of life on Earth. Its influence is marked not just by the research it has motivated, but also by the theories and applications it has spawned. Many present day strategies in reserve design, landscape ecology and metapopulation theory draw significant portions of their intellectual capital from ETIB. Indeed, it is hard to imagine what the modern face of ecology, biogeography or conservation biology would be like without ETIB.

The Equilibrium Theory of Island Biogeography was first articulated by Robert H. MacArthur and Edward O. Wilson in a short journal article in 1963 and later expanded to a monograph in 1967. It was this second publication that garnered the attention necessary for this theory to develop into the mainstream research juggernaut that we know today. By the early 1970s many of the best, and perhaps prophetically, most influential researchers in the fields of ecology and conservation biology were actively engaged in testing or expanding ETIB (e.g., Brown 1971; Terborgh 1973; Simberloff 1974; Diamond 1975). Through the 1980s and 90s theories influenced by ETIB became predominant in the literature. Today, the topic of ETIB is still actively investigated (e.g. Heaney 2007; Whitaker et al. 2008) and its continuing legacy is pervasive in recent literature. Indeed, MacArthur and Wilson's work (1967) was the focus of a recent symposium at Harvard University, and a subsequent edited volume, which celebrates and examines the influence of ETIB over four decades (Losos and Ricklefs 2009).

MacArthur and Wilson (1967) aimed to shift biogeography from a descriptive natural history phase to "an equally interesting experimental and theoretical phase." (p. 181). Specifically they wanted to construct a "theory of biogeography at the species level" (p. 5) – one that deemphasized the traditional focus of biogeography on the composition of biotas, the distribution of higher taxa and the role of geological change. They believed that biogeography could "be reformulated in terms of the first principles of population ecology and genetics." (p. 183) This perception suited their conception of how theories should be developed, as they suggested that "A theory attempts to identify the factors that determine a class of phenomena and to state the permissible relationships among the factors as a set of verifiable propositions." (*p*. 5)

In this chapter we outline the conceptual basis, predictions, impact and future of the study of the Equilibrium Theory of Island Biogeography. ETIB represents the best known and most investigated portion of the total body of theory described by MacArthur and Wilson in their monograph (1967). Many other topics were also discussed at length in their book, e.g., the strategy of colonization, invasibility and the variable niche, etc. We restrict our focus in this chapter, however, to ETIB itself. Further, we do not attempt to provide a review of the work that preceded ETIB and laid its conceptual ground work, e.g., by Dammerman (1948), Darlington (1957), Wilson (1961) and others, as thorough reviews on this topic are available elsewhere (see Whittaker and Fernandez-Palacios 2007; Lomolino et al. 2009; Lomolino and Brown, in review). Here, we begin by describing the domain and propositions of ETIB. We then explore how these propositions are built from deductive reasoning and fundamental principles in ecology and physical sciences. Next we examine the degree to which evidence has supported or refuted the theory. We also consider how ETIB relates to other models of species diversity and how ETIB has influenced the development of other theories in ecology. We consider the case-study of anthropogenic invasions of species to islands with respect to ETIB. Finally, we consider how ETIB can be improved and what areas of research ETIB can productively lead us towards in the future.

DOMAIN

The Equilibrium Theory of Island Biogeography provides an explanation for variation in patterns of species diversity across space and over time for islands and insular habitats. MacArthur and Wilson (1967) believed that insularity is "*a universal feature of biogeography*" (p. 3) and should be applicable in a broad array of habitats. They suggested that "*Many of the principles graphically displayed in the Galapagos Islands and other remote archipelagos apply in lesser or greater degree to all natural habitats. Consider, for example, the insular nature of streams, caves, gallery forest, tide pools, taiga as it breaks up into tundra, and tundra as it breaks up into taiga.*" (p. 3-4). Because 'insularity' is not a discrete trait of habitats, the degree to which a given habitat should fall within the domain of ETIB should be related to its degree of isolation from other suitable habitats, i.e. the degree to which the matrix that fills the space between focal-

habitats impedes dispersal by those species that occupy the focal-habitats. Further, whether or not islands or insular habitats are natural or anthropogenic should not influence whether they fall within the domain of ETIB. All areas that are discreetly partitioned from a surrounding matrix should have characteristic patterns of species richness and species turnover that are coincident with the size and isolation of those partitioned areas. MacArthur and Wilson (1967) suggested therefore that "*The same principles apply, and will apply to an accelerating extent in the future, to formerly continuous natural habitats now being broken up by the encroachment of civilization*". (p. 4). Consequently, the domain of ETIB can be envisioned to include much of the natural and human-modified habitats that occur globally.

The spatial and temporal extent of the domain of ETIB, however, is not universally agreed to be so extensive. Many authors (e.g., Whittaker 1998, 2000; Whittaker and Fernandez-Palacios 2007; Lomolino et al. 2009) have suggested that the spatial extent of ETIB should be restricted to a much narrower set of conditions than originally envisioned. For instance, Lomolino et al. (2009) argue that only islands at relatively intermediate levels of isolation and at intermediate sizes should be well-characterized by dynamic equilibrial processes described in ETIB; for example, Lomolino et al. (2009) posit that on very small, near islands that species diversity should be characterized primarily by stochastic processes and little influenced by equilibrial ones. We discuss this perceived contraction of the applicable domain of ETIB in the section on 'Evaluating ETIB'. Additionally, disagreement arises over the temporal domain of ETIB. In its most basic form ETIB applies to patterns of species diversity and turnover over relatively short, ecological-time scales. Many authors consider ETIB solely in this temporal context (see discussion in Whittaker and Fernandez-Palacios 2007). However, one can also consider the longer-term processes that MacArthur and Wilson (1967) postulated would impact equilibria over geological and evolutionary time-scales.

FUNDAMENTAL PRINCIPLES UNDERLYING THE PROPOSITIONS OF ETIB

The Equilibrium Theory of Island Biogeography, as originally envisioned by MacArthur and Wilson (1963, 1967), can be partitioned into a basic model that operates over 'ecological time' and a more inclusive model that operates over 'evolutionary time'. It is the basic model, over ecological time, that has been influential and remains well known. These two models are discussed separately below.

Principles underlying the ecological model of ETIB

There are seven propositions of the basic, ecological model of ETIB (Table 10.1). Proposition 1 is derived from deductive reasoning. The rate of immigration (i.e., arrival of new species) to an island must decrease as more and more of the species that could potentially arrive have done so. Further, this rate must reach zero once the entire pool of species that could colonize are present. If species are considered as equivalent units, i.e. with equal probabilities of arrival, then this rate should decrease linearly with increasing richness (Figure 10.1A). This proposition is refined, however, by MacArthur and Wilson, who invoke differences among species, wherein some species are more likely to arrive to an island per unit of time than others. This results in the classic downward 'bow' in the immigration curve depicted in ETIB (Figure 10.1B). Note, however, that as long as the immigration curve decreases monotonically the precise form of this curve will not change the qualitative predictions of ETIB.

Propositions 2 and 3 are in accord with the physical process of diffusion. With respect to Proposition 2, for example, objects diffusing from a source will reach a near object more frequently then they reach a far object. This means that rates of immigration should be higher on near islands than far islands, as species by chance should more frequently encounter near islands (Figure 10.1B). Nevertheless, both far and near islands could be colonized by all species in the pool of potential colonists given sufficient time, such that the rate of immigration for near and far islands both reach zero once all species from the pool are present (Fig. 1B). Proposition 3 is based on the same process of diffusion, but with regard to target size – such that arrival of colonists will be more frequent to a large island than to a small one. This proposition was described by MacArthur and Wilson (1963); in the monograph (1967), it was excluded from their introductory description of ETIB, but included in their description of stepping stone islands, and thus retained within the larger suite of factors postulated to influence rates of immigration.

Propositions 4 and 5 are derived from deductive reasoning and knowledge of natural history (particularly the study of population ecology). If all species have an equal probability of going extinct per unit of time, and this process is independent of the total number of species, then the rate of extinction will increase linearly with an increasing number of species (Figure 10.1A). More realistically, however, species may vary in their individual probability of extinction, which may impact the rate of extinction for any given number of species. Further, as the number of

species that are using a given set of resources increases, the average number of individuals per species must decrease, so long as resources are being fully (or nearly fully) used; because smaller populations are more likely to go extinct than larger populations the rate of extinction may not be independent of the total number of species. Additionally, increasing the number of species may increase the probability that species will interfere with each other. These factors may result in an extinction rate with a downward 'bow' from a straight line, i.e., one that increases non-linearly with an increasing number of species (Figure 10.1B). Note, however, that MacArthur and Wilson (1963, 1967) do not incorporate concepts of positive species interactions, such as facilitation and mutualism, into their rationale for extinction rates (but see Wilson 1969); doing this could alter the relative position of the extinction curve either upwards or downwards, but should not change its basic shape (Wilson 1969). As with the immigration rate, however, so long as the extinction curve increases monotonically the precise form of this curve will not change the qualitative predictions of ETIB. Finally, for any given number of species, a larger island should have a lower rate of extinction, since average population sizes would be larger and hence these species would be less likely to go extinct (Figure 10.1B).

Propositions 6 and 7 are in accord with equilibrium processes observed in chemistry, physics and population biology. For example, the total number of individuals in a closed population will be determined by an equilibrium between opposing rates of births and deaths. By altering either of these rates the total number of species in the population can be altered. MacArthur and Wilson (1967) similarly reasoned that the number of species supported on an island will depend on an equilibrium between the rates of immigration and extinction (Figure 10.1A). By varying the rates of immigration and extinction, the resulting equilibrium number of species can be shifted (Figure 10.1B); these shifts in equilibrium points provide an explanation for species-area and species-isolation relationships. Rates of species turnover are also derived from equilibrium processes, such that islands with the highest rates of immigration and extinction even though species number remains constant (Figure 10.1B). Although MacArthur and Wilson (1967) predicted that turnover rates should vary with island characteristics, the graphical depiction (as shown in Figure 10.1B) appeared in later publications (e.g. Simberloff 1974).

Principles underlying the evolutionary model of ETIB

As suggested above, not all authors agree that evolutionary considerations should be included as part of ETIB. Here, however, we discuss those propositions described by MacArthur and Wilson (1967) that are expected to directly influence the predictions of ETIB over evolutionary timescales (Table 1). The principles underlying evolutionary components of ETIB are based on deductive reasoning (Proposition 8), observed patterns in the geographical distribution of endemic species (Propositions 9 and 10), and principles of evolutionary biology (Proposition 11). Deductively, it is clear that species could be added to an island by either a process of immigration or by *in situ* speciation events (Proposition 8). Geographic patterns in the distribution of endemic species, particular of ones believed to have evolved unique species status on the island or archipelago in question, show that the importance of speciation increases with isolation (Mayr 1965). MacArthur and Wilson (1967) reasoned that speciation will be most important at the outward margin of the zone that species are capable of dispersing to, presumably because subsequent gene flow from mainland sources would be quite low; this should result in a "radiation zone", where speciation is more likely to occur (Proposition 9). Further, because taxonomic groups vary in their ability to disperse, the location of radiation zones should be taxon specific (Proposition 10). Finally, MacArthur and Wilson (1967) believe that "an equilibrium can be defined in either ecological or evolutionary time." (p. 176). Consequently, an ecologically observed equilibrium is really a "quasi-equilibrium" (p. 176) that can be increased over evolutionary time (Proposition 11). They expect this to occur because of evolutionary adaptation among species and to the environment over time (MacArthur and Wilson 1967, Wilson 1969).

EVALUATING ETIB

Overall, and despite its iconic nature, ETIB has been difficult to test. Because the equilibrium model was designed to provide an explanation for diversity increases with increasing island area and decreases with increasing isolation, confirmation of these patterns is not strong evidence in support of ETIB, particularly as other models of diversity (see section below) share these predictions. Nevertheless, it is worth noting that these patterns (even with 40 years more data) often do bear out as expected; for example, a recent study of the vascular floras of 488 islands worldwide found that area alone predicted 66 percent of the variation in species richness, and that the next most important predictors were isolation, temperature and precipitation (Kreft et al. 2008). A thorough evaluation of ETIB, however, and the hundreds of papers that consider its

predictions are beyond the scope of this chapter; see Whittaker and Fernandez-Palacios (2007) for an overall review. Here, we provide a brief sketch of how the predictions of ETIB have fared in the literature, in particular with respect to species richness and turnover.

There is mixed support for several of the specific predictions made by ETIB for species diversity. First, ETIB predicts that the slope of species-area relationships will increase with increasing isolation. This has proven true in some comparisons, particularly between classes of islands, e.g., vascular plants on landbridge versus oceanic islands (Sax and Gaines 2006), but has not proven true in many other cases (Whittaker and Fernandez-Palacios 2007). Second, ETIB predicts that an integration of immigration and extinction curves will produce a 'colonization curve', representing the total accumulation of species on an island, that should asymptote over time. Evidence that is arguably consistent with this has been observed on small, experimental islands (e.g., Simberlof and Wilson 1970), but few studies are available for larger islands. Some of the best evidence on larger islands comes from Krakatau - but even in this case evidence of an asymptote in species numbers is difficult to evaluate, because the time-scale of an expected asymptote is not well defined by ETIB. Consequently, it is unclear how to evaluate the shallowly increasing slope of bird richness or the more sharply increasing slope of plant and butterfly richness (Whittaker 1998). Third, ETIB predicts that declines in area should drive species losses in the smaller more isolated area of habitat remaining. This process is referred to as 'relaxation' in species numbers. Evidence consistent with this expectation does exist, e.g., for the loss of species in Singapore (e.g., Brook et al. 2003). In many cases, however, this expectation is not well supported, e.g. many oceanic islands have lost much of their habitable area to agricultural and urban areas, yet few plant species have gone extinct (Sax et al. 2002). Here too, however, this evidence is difficult to interpret, because the timescales necessary for relaxation to occur are poorly defined by ETIB; it is conceivable, for example, that extinctions for certain taxa may only manifest after many hundreds of years (Sax and Gaines 2008).

The primary focus of testing and falsifying the ecological model of ETIB has been in evaluating its predictions of species turnover. Here too, the evidence in support of ETIB has been mixed at best, with a few studies in apparent (but debated) support, but many other studies that don't or only partially support the theory's predictions. Two of the best known studies posited to support ETIB are the study of insect immigration following the experimental defaunation of small mangrove islets (e.g., Simberloff and Wilson 1969) and the turnover of land birds on the Channel Islands of California (Diamond 1969). Both showed turnover of species with relatively stable equilibria in species numbers. Evidence in support of ETIB from Simberloff's work has been questioned because of the difficulty in distinguishing transient species from actual immigrants and for being done on small, experimental systems that may not scale up to larger islands (Whittaker 1998). Diamond's work has been critiqued because estimates of turnover may have been inflated by incomplete census data and by anthropogenic effects (Lynch and Johnson 1974). Mixed support for ETIB can be seen in wide-variety of other studies. For instance, Cody (2006) in a 25-year study of nearly 200 continental islands in the Barkley Sound, British Columbia found that rates of immigration and extinction on islands were highly correlated, such that stable equilibria in species numbers were maintained (in support of ETIB). Yet rates of immigration and extinction were highest on large islands, regardless of isolation. As a result, turnover was highest on large islands - in contrast to predictions of ETIB; although in this case because the absolute difference in isolation among these islands is relatively small it is possible to conclude that these findings are not in conflict with ETIB. Finally, there are cases where the predictions of ETIB do not hold well at all, e.g., for orb-weaving spiders in the Bahamas (Toft and Schoener 1983).

Overall, there are several inherent difficulties in testing turnover rates. First, some of the best evidence comes from small experimental systems, which present challenges in generalizing to larger scales. Second, studies performed at larger spatial scales are generally performed over relatively short periods of time, but see Cody (2006), making it difficult to generalize to longer timer periods. Third, complete and comprehensive census data are rarely available, making it difficult to evaluate actual turnover rates (Whittaker and Fernandez-Palacios 2007). Fourth, the definitions of immigration, colonization and extinction are poorly articulated, in that it is unclear what determines whether a species is established on an island, which in turn influences whether a species is classified as a failed immigrant or instead as an established species that went extinct. Further complicating these evaluations are differences in how ETIB is defined. Some authors consider only the ecological model of ETIB, while other authors consider the evolutionary model and other refinements discussed in the 1967 monograph. Consequently, the same evidence posited against the equilibrium model by one individual, might be held in support of the broader, more inclusive view of ETIB by another (Whittaker 2000). Given these many complications,

Whittaker (1998) has suggested that it is difficult to find any convincing, unequivocal evidence of turnover that matches the predictions of ETIB.

Given the many difficulties with testing ETIB, one might ask why it has been so influential and why it remains such an active focus for research. The answer has much to do with the compelling simplicity of a model that offers to explain complex patterns in species richness and turnover. As Heaney (2000), Lomolino (2000a) and others have suggested, it is ETIB's heuristic features that make it so powerful. It provides a frame-of-reference for observations about the natural world, particularly ones that differ from expectations of ETIB. Undoubtedly, its influence is further heightened by two features. First, that no alternative model has gained primacy in the literature. Second, that ETIB has influenced the development of many other theories in ecology and conservation biology.

CONTRASTING ETIB WITH OTHER MODELS OF SPECIES DIVERSITY

Since a central goal of ecology and biogeography is to explain the distribution and diversity of life, it is not surprising that many alternative theories exist for patterns of species diversity. Some of the most influential include: neutral theory (Hubbell 2001), stochastic niche theory (Tilman 2004), and species-energy theory (Wright 1983). Other 'unconsolidated' theories also exist, such as the idea that diversity patterns on islands are largely driven by non-equilibrial aspects of dispersal limitation. We do not attempt to provide an exhaustive review of this topic, which would include discussion of longstanding models by Lack (1947), as well as more recent work by Heaney (2000), Lomolino (2000b), O'Brien et al. (2000) and others. Here, we highlight the importance of dispersal limitation in influencing diversity patterns, along with just two of the more influential, codified theories so as to provide a context for better understanding ETIB.

Dispersal limitation and a non-equilibrial view of island diversity

The ideas comprising a non-equilibrial view of island diversity have not been codified into a named hypothesis or theory. Nevertheless, this view is prominent in the literature. It involves a similar and overlapping set of ideas often referred to with slightly different language: 'undersaturated' (Lawlor 1986), 'non-equilibrial' (Whittaker 1998), 'not saturated' (Sax and Gaines 2008), etc. Put simply, this view holds that diversity is not a consequence of an equilibrium between opposing forces. Like ETIB, this view posits that diversity could be

increased if dispersal limitations were removed or if rates of speciation were increased, but unlike ETIB, it does not posit that these factors are necessarily opposed, at least in part, by increases in extinction. In principle, distinguishing this view from ETIB is difficult in many cases because the time-scale of the key processes are poorly defined. So, for example, if immigration rates are rapidly increased by anthropogenic means to an island, how long will it take to see increases in extinction rates, as predicted by ETIB? Will some of the increases in diversity be transient and offset by future extinctions, or will the increases in diversity be persistent? See the section, below, on anthropogenic invasions for a more complete discussion of this issue. In other cases, however, it appears certain that a dispersal-limitation, non-equilibrium model is an appropriate way to understand patterns of diversity. For example, natural rates of immigration for mammals on remote oceanic islands appear to be close to zero (e.g., Lawlor 1986). Indeed, a non-equilibrial explanation is consistent with observed patterns for many dispersal-limited groups: non-volant mammals, freshwater fishes, amphibians, and others.

SPECIES-ENERGY THEORY

Species-Energy Theory as proposed by Wright (1983) is a direct derivation from ETIB, although the precursors to this theory include much older thoughts on the importance of energy in determining numbers of species (e.g., Forster 1778). In Wright's (1983) model the principal difference with ETIB is the substitution of 'energy' for 'size' of an island. In relation to ETIB, this means that extinction rates would be set by available energy instead of island size. Similarly, Wright substituted energy-area relationships for species-area relationships. He described area with Actual Evapotranspiration (AET). This metric of 'energy' has the advantage of being readily measurable for most places. With these substitutions, Wright found that he increased the amount of variation typically explained by species-area relationships (SARs). He also suggested that this would allow the framework of ETIB and SARs to be applied across disparate regions of the world (e.g., in temperate and tropical regions), while using a single explanatory framework. Wright's theory has been adopted and investigated by many other researchers (e.g., Turner et al. 1988). The most influential papers on species-energy theory, however, appear to be motivated independently from Wright's work. These papers by Currie and Paquin (1987) and Currie (1991) examine the relationship of species diversity of plant and animal groups on continental areas. However, Wylie and Currie (1993) apply these same ideas to mammals on land bridge islands,

and like Wright (1983), who examined plants, find that available energy is an excellent predictor of species richness. Similarly, Kalmar and Currie (2006), who explicitly set out to test ETIB and species-energy theory, find that area, climate and isolation are all important predictors of bird diversity on islands worldwide, but in contrast to predictions of ETIB, isolation does not influence the slope of the species-area relationship (SAR); Whittaker (2006) suggests, however, that the failure to find an affect of isolation on the slope of SARs may be due to mixing islands from different climatic zones together into a single analysis. Much research is continuing along these lines currently (see Fox et al. Chapter 13), and this work offers the potential for major advances in our ability to predict patterns of species diversity.

STOCHASTIC NICHE THEORY

Stochastic Niche Theory (Tilman 2004) is a modern reformulation of resource use among species that provides an explanation for patterns of species invasion or colonization and observed levels of species richness. It is similar to ETIB in that its core model effectively treats species as equal entities, while it also provides a larger construct in which species differences could allow for refinements of the model (Stachowitcz and Tilman 2005). The model differs from ETIB in having 'priority' effects in which species that arrive first are most likely to retain their place in a community (or on an island). So, instead of having a dynamic equilibrium set by opposing rates of immigration and extinction (in which the addition of a species should result in the loss of an existing one), Stochastic Niche Theory posits that species that colonize should continue to be added so long as there are sufficient available resources. Once those resources start to become limiting, then the probability of adding additional species should decline in proportion to species richness; at some point of resource use the total number of species should be very difficult to increase – this level of richness may appear as a saturation point, even though it would be generated by non-equilibrium processes. This model has been relatively influential to date, but more time will be needed to judge its eventual impact on the field.

THE RELATION OF ETIB TO OTHER THEORIES IN ECOLOGY

First and foremost, ETIB is an extension of basic population biology theory. Population biology theory treats all individuals within a population as equals, with respect to growth rates, carrying capacity, etc. This is a simplifying assumption, as clearly not all individuals are truly equal –

some are larger, some more fecund, etc. Nevertheless, this simplifying assumption allows for models explaining general patterns and trends among individuals within a species to be understood and predicted (see Hastings Chapter 6). ETIB borrows concepts from population biology theory that are applied to individuals and instead applies them to species. For instance, ETIB treats species as largely equivalent units. Further, instead of a 'carrying capacity' for individuals set by opposing rates of births and deaths, ETIB posits a saturation point in species numbers set by opposing rates of immigration and extinction. Like population biology theory, wherein the carrying capacity for individuals can be increased, so to can the saturation point for species be increased in ETIB, whenever changes in equilibrium processes are changed. Certainly, ETIB goes beyond population biology theory in many ways, but its similarities are readily apparent.

ETIB is also intricately related to many other theories in ecology. Indeed, the ultimate legacy and influence of ETIB is perhaps not found in the model's predictive power, but instead in the great diversity of theories that ETIB has influenced or helped to generate. The influence of ETIB on all of these models illustrates the power of clearly articulating a simple heuristic theory, as this provides the foundation for further advances. In many cases, these models and theories have perhaps superseded the influence of ETIB itself. In the section above, we have described the relationship between ETIB and species-energy theory, stochastic niche theory and non-equilibrium views, but there are several other prominent theories that have been influenced by ETIB. We highlight four of these below.

ETIB has influenced the development of metapopulation models. Although the origin or metapopulation models predate ETIB (see discussion in Hanski 1991) the modern conception of these ideas was developed by Richard Levins (1969, 1970) after the development of ETIB. Levins' codified the definition of a metapopulation as "*a population of populations which go extinct locally and recolonize*". Consequently, at its core, a metapopulation model borrows basic ideas from ETIB of immigration and extinction dynamics affecting a network (or archipelago) of local patches. This model has had broad and wide-ranging influences on the fields of ecology and conservation biology, influencing subfields like population ecology (Freckleton and Watkinson 2000) and metacommunity theory (see Leibold Chapter 8).

ETIB has also influenced the scientific underpinnings for reserve design. This began with a series of publications that followed the publication of MacArthur and Wilson's work, e.g.

Terborgh (1974), Wilson and Willis (1975), but was strongly influenced by the publication of Jared Diamond's (1975) seminal work on reserve selection. Diamond's work built on MacArthur and Wilson's supposition that ETIB should apply to habitat fragments; he extended this idea to consider the pros and cons of alternative reserve designs. This resulted in the famous SLOSS (Single Large or Several Small) reserve design debate (Gottelli 2004), but also formed the basis for consideration of issues in reserve design and planning that continue to this day (e.g., Pressey et al. 2007). Some authors believe, however, that while ETIB has been important in the development of this field that current work transcends the boundaries and explanations originally proposed by the theory (e.g., Laurance 2008).

ETIB strongly influenced the development and codification of the field of "macroecology" - a term coined by Brown and Maurer (1989). Macroecology examines the statistical properties between the dynamics and interactions of species populations (Brown 1995). It posits that advances in our understanding of the natural world can be advanced by observations of general trends, and a search for emergent properties, particularly when comparing species within given taxonomic groups (such as birds or plants). Consequently, macroecology as an 'approach' has proven to be a powerful tool, one that extends well beyond the domain of ETIB, to consider patterns of body size, range size, allometric scaling, speciation and other critical concepts in ecology and evolutionary biology (e.g., Brown 1995; Enquist et al. 1998; Blackburn and Gaston 2003; Allen et al. 2006).

ETIB has influenced the development of "neutral theory", as outlined by Stephen Hubbell (2001). Like ETIB, this theory is inherently a dispersal-assembly model (built on concepts of immigration, speciation and extinction). Perhaps even more than ETIB itself, Neutral Theory extends the concept of symmetry (i.e. equivalency) between species as a simplifying assumption that allows important patterns of species distribution and abundance to be examined against a 'null' hypothesis. Like ETIB, this model has generated great debate and in so doing it appears to have pushed the field forward. It is too early to judge what its ultimate influence will be, but it has the potential to have a strong, multi-decadal trajectory, not unlike that of ETIB.

A CASE STUDY OF ETIB: ANTHROPOGENIC SPECIES INVASIONS OF ISLANDS Human-facilitated colonization events, i.e. species invasions, have provided researchers with a novel source of insight into ecological, evolutionary and biogeographical questions since at least the 19th Century (Sax et al. 2007). Darwin (1859), Grinnell (1919), Baker and Stebbins (1965), as well as many recent investigators have viewed invasions as a set of 'unplanned' experiments. For instance, Huey et al. (2000) used invasions of fruit flies in North and South America to examine the speed and predictability of evolution of clines in wing size across geographic gradients. Similarly, Bruno et al. (2005) used invasions as a source of evidence to examine the relative roles of competition, predation and facilitation in newly-assembled species associations. It is, perhaps, not surprising that researchers have also used invasions to better understand patterns of species richness and turnover on islands (e.g., Sax et al. 2002; Sax and Gaines 2008).

MacArthur and Wilson (1967) were not silent on the topic of species invasions, but instead made a set of observations and predictions regarding them. First, they conjectured that species addition (whether through natural or human-facilitated immigration) must lead to a reduction in the population size of existing species "...unless, as in the case of the more recently introduced starling and kiskadee on Bermuda, the species are able to tap a previously unused food supply." (p. 97). Second, they reasoned that "By increasing the immigration rate, an impoverished biota can be changed into a richer one, yet without altering its equilibrial condition in the end. It will merely shift from one saturated state to another." (p. 176). This suggests that increased introduction rates of species (whether caused by humans or not) could lead to increases in the stable equilibrium number of species that inhabit islands. Third, they predicted that such increases would not be permanent if immigration rates were subsequently reduced. Consequently, they suggest that "If the new inpouring of immigrants were to be held constant, the number of Hawaiian bird species – native plus introduced – would move to a new, much higher equilibrium level....If, on the other hand, all further importations were strictly forbidden so that the immigration rate returned to the old, natural level, the number of species might gradually decline to a third equilibrium not radically different from the pre-European level." (p. 177).

Given MacArthur and Wilson's (1967) predictions, as well as the general usefulness of invasions for testing theory, it seems worthwhile to consider the implications of species invasions for the evaluation of ETIB and reciprocally for understanding the ultimate impact that invasions may have on native taxa. ETIB, as described by MacArthur and Wilson (1967), needs to be modified only slightly to consider human-mediated invasions. One change to classical ETIB that is needed occurs because human activities can increase not just rates of immigration, but also the total potential number of colonists, such that the pool of potential colonists is increased (Fig. 2). This occurs because humans are transporting species not just from the source pool that originally stocked a given island, but also from various other parts of the world, where species were previously unable (or at least extremely unlikely) to be drawn from by natural processes of dispersal. Some authors, such as Whittaker and Fernandez-Palacios (2007), have suggested that human alterations of the environment might also have increased the total carrying capacity for species on islands, e.g., by removing forest cover and creating room for smaller herbaceous species to invade. This is certainly conceivable, in which case this could be modeled, with respect to ETIB, by sharply lowering the extinction curve – as the effective 'size' of an island would be larger. Here we review the evidence of species invasions and turnover, with respect to ETIB, on oceanic islands in four taxonomic groups: freshwater fishes, vascular plants, non-volant mammals and land birds.

Freshwater fishes and vascular plants

These two groups show qualitatively similar patterns of invasion and species turnover on islands. Both groups have seen dramatic increases in total number of species established on oceanic islands, with many successful invasions and few extinctions (Sax et al. 2002, Sax and Gaines 2003). This has occurred because many non-native species have become established, while few native species have gone extinct. In freshwater fishes, the most extreme example is in Hawaii, where all five native species have persisted in spite of the successful invasion of 40 non-native species – such that total species number has increased by 800% (Eldredge and Miller 1995). Vascular plants have also experienced a large increase in net richness of species, with an extremely consistent near-doubling in species number across oceanic islands of the world (Sax and Gaines 2008). In New Zealand, more than 2000 non-native plants have become established, largely matching the slightly more than 2000 native plant species extant, with only a handful of native extinctions. Further, in New Zealand there is a sufficient fossil record to ascertain that a large fraction of the flora has not recently gone extinct, i.e. that the few recorded extinctions are not an artifact of incomplete data (Sax et al. 2002).

With respect to ETIB, for both fishes and plants, these patterns are best fit by an increased immigration rate (and pool size), as well as by an 'observed' extinction rate that is
nearly-flat, but increasing very slightly with increased richness (Fig. 2A). The relatively flat extinction rate could be due to an overall increase in the carrying capacity of islands caused by anthropogenically-based increases in habitat heterogeneity. Alternatively, the few extinctions observed to date could conceivably be due to long time-lags. Consequently, if there are significant time-lags associated with 'relaxation', i.e., change from a supersaturated back to a saturated state, then it is possible that the ultimate extinction rate could be much higher (Fig. 2A). If this were the case, the current increases in diversity would be transient and future species richness should decline as a result of processes that have already been set in place. This relaxation would not be expected to occur all the way back to pre-disturbance levels, unless the rate of immigration were also reduced. However, a reduction in immigration rate in the near-term seems extremely unlikely, at least for plants, because a large pool of potential colonists have already been transported to many islands. For example, in New Zealand there are more than 20,000 non-native plant species, which have not yet become established, that currently reside in gardens around the country (Sax and Gaines 2008).

Although the patterns observed with fishes and plants can be viewed within the context of ETIB, they can also be considered within the context of a non-equilibrium state in which species richness on islands has been undersaturated (*sensu* Lawlor 1986) because of dispersal limitation. The extreme dispersal limitation experienced by freshwater fishes in crossing saltwater barriers is consistent with this alternative explanation for observed patterns of species richness. Similarly, the observation that the number of naturalized plant species has been increasing linearly over the past few hundred years, with no signs of an asymptote on more than a half-dozen islands that have adequate historical records to make such determinations, further conflicts with expectations of ETIB (Sax and Gaines 2008). Additional work will be necessary to fit the best theoretical framework for interpreting anthropogenic patterns of invasion in fishes and plants, as well as future extinctions in native species.

Non-volant mammals and land birds

Mammals and birds have shown very different patterns of invasion on islands. Non-volant mammals have historically been absent from most oceanic islands because of dispersal limitation; their richness has increased greatly with human introductions (Lever 1985). Consequently mammals appear to fit a non-equilibrium model on oceanic islands. In contrast,

land birds, which are effective dispersers, are species rich on many oceanic islands. Their pattern of change in richness following human introductions and habitat disturbance has been more complex. Richness of land birds initially decreased following colonization of islands by humans and their mammalian commensals (mice, rats and cats), with large numbers of bird extinctions. Human introduction of birds, along with natural colonization following habitat transformation (e.g. conversion of forest to pasture) has increased bird richness back to levels that existed prior to human colonization, i.e. prior to colonization by either aboriginal or European peoples (Sax et al. 2002). This has occurred in a highly non-random manner such that the number of birds driven extinct on oceanic islands is closely correlated with the number that have subsequently become established ($R^2 = 0.78$, p < 0.001). This pattern of change in birds can be envisioned in a ETIB context if an increased immigration rate has been matched by an increased extinction rate (Fig. 2B). An increased extinction rate for birds is conceivable if introduced mammals, which can compete for resources with birds or prey upon them, have reduced the carrying capacity for birds on islands, i.e. if the 'size' of islands for birds has affectively been reduced. Another way that these data could be viewed to be consistent with ETIB is if we consider birds and mammals together, as a single group. Given the many mammal-like roles that birds have filled on oceanic islands that were historically lacking mammals, such an amalgamation may be reasonable, although certainly unconventional. In this case, when viewed collectively as a single group, the net change in richness and turnover (of birds and mammals) fits well with an ETIB model (Fig. 2C), as net richness and turnover have increased. Additional work will be needed to ascertain how best to encapsulate future changes in diversity on islands for birds and mammals.

MAPPING A ROUTE FORWARD TO IMPROVE ETIB

The Equilibrium Theory of Island Biogeography, as originally conceived, is not sufficient to provide a universal model for understanding patterns of species richness or turnover on islands (Whittaker 1998; Lomolino 2000a; Heaney 2007). Whittaker (1998, 2000), Lomolino et al. (2009) and others have suggested that ETIB, as a dynamic equilibrium model, may be sufficient only within a subset of the larger range of conditions that exist in nature, from equilibrial to disequilibrial and from static to dynamic systems. Nevertheless, as a heuristic device ETIB has been and continues to be extremely important in organizing thought on patterns of species diversity and in providing a framework against which to consider other models of diversity. The

question remains whether modifications and improvements on the basic ETIB model can help to provide a more robust framework for evaluating and predicting patterns of species diversity or whether sharper departures from ETIB are needed. Here we outline three areas of research that may help to determine the most productive roads forward.

Integrating ecological and evolutionary processes

Ecological and evolutionary process are both considered within MacArthur and Wilson's work (1963, 1967) and by Wilson (1969). This was done in two ways. First, MacArthur and Wilson (1963, 1967) suggested that species addition to islands could occur from immigration and speciation. As discussed above, they believed that the latter should only be important in 'radiation zones' far from mainland sources of species. Second, they stated that "*In evolutionary time, a gradual increase in the species number is permitted through adaptation of the immigrant species to each other and to the local environment over a relatively long period of time.*" (p.

176). This suggests that an ecologically-based, immigration-driven equilibrium can be increased through evolutionary processes over time. This second view was further developed by Wilson (1969), who argued that the relative influence of this adaptation over evolutionary time could account for the more species-rich assemblage of native ants versus non-native ants on islands in Polynesia. Neither MacArthur and Wilson (1963, 1967) or Wilson (1969) provide quantitative estimates of how long these evolutionary processes may take to occur. Further, through a process of 'domain contraction' the evolutionary components of ETIB were largely set aside in the 1970s and 1980s (Lomolino et al. 2009).

A new paradigm appears to be emerging (e.g., Lomolino 2000a; Heaney 2007; Whittaker and Fernandez-Palacios 2007; Lomolino et al. 2009) that an improved theory of island biogeography must more explicitly consider the simultaneous influence of speciation, immigration and extinction, as well as ecological interactions and differences among species in all these features. Such a paradigm would more explicitly recognize the important role that adaptive radiation plays in increasing species diversity on islands (Heaney 2007). Over the last 40 years there have been important advances in our understanding of speciation on islands (e.g., Funk and Wagner 1995; Gillespie 2004), but these improvements have remained divorced from integrated ecological and evolutionary models of species diversity. Consequently, there is much need for work on this topic, which offers the potential to more robustly describe the distribution of life on Earth.

Incorporating island ontogeny

ETIB considers few characteristics of islands beyond their size and isolation. This simplifying step has provided the basis for model development, but is perhaps not sufficient for the integration of ecological and evolutionary processes described above. Recently, Whittaker et al. (2008) have suggested that it may be important to consider island 'ontogeny', i.e. the life-history of an island itself. Consider, for example, that a volcanic island's size, topographic complexity, ability to support life, and potentially its likelihood of promoting speciation may all change over geological time as the island forms out of the ocean, grows larger and higher, then begins to shrink, lose elevation and eventually become eroded below the surface of the sea. Such a model of island ontogeny explicitly considers the possibility of within-island allopatry, i.e. allopatry that occurs within a topographically complex island. This is distinct from MacArthur and Wilson's (1967) description of allopatry, in which only among-island allopatry is considered. Models that incorporate island ontogeny would also fit well within the framework suggested by Heaney (2000) in which species diversity was regularly within a dynamic state of disequilibrium, always a few steps behind constantly changing geological or geographical circumstances. This promising avenue of research offers one important inroad to improving ETIB.

Borrowing aspects of ETIB to study mainlands

A variety of recent work on mainlands has shown that species richness can often remain remarkably stable over time or across space despite large changes in species turnover. In contrast to ETIB, these mainland environments occur in habitats that are artificially delimited, i.e. they are not marked by discrete physical boundaries like islands or insular habitats. For example, Brown et al. (2001) have shown within delimited areas that birds in Michigan, mammals in Arizona, and trees in Europe have retained relatively stable numbers of species through time despite large changes in species turnover. Similarly, across space, Sax et al. (2005) have shown that analogous communities on mainlands that are dominated by non-native species often support very similar numbers of species as native communities, despite having very different species compositions. Brown et al. (2001) have suggested that equilibrial numbers of species may be maintained, despite turnover in species composition, by interaction between a carrying capacity for species set by local ecological conditions and by opposing rates of immigration and extinction. This is intriguing, in part because Rosenzweig (1995) and others have suggested that immigration pressure will generally not be limiting in mainland environments. Ironically, if the model proposed by Brown et al. (2001) bears out then it may suggest that cases of dynamic turnover in species identity with relatively stable numbers of species may be most common in places where immigration pressures are not limiting, i.e., on mainlands and not on islands. Further work is needed to better explore this possibility.

IMPLICATIONS OF FURTHER EXPLORATION OF ETIB FOR ECOLOGY, EVOLUTION AND CONSERVATION

ETIB has served as a powerful heuristic tool for advancing our understanding patterns of diversity in time and space. Recent attempts to test and expand the theory have raised two issues that are particularly pressing for conservation biology, but also for our continued advance of ecological and evolutionary theory. First, we need to better characterize the rates at which species extinctions are likely to occur once processes have been set in motion that may 'commit' species to eventual demise. The time-lags and 'extinction debt' involved in such extinction processes are still poorly explored and in need of much attention (Tilman et al. 1994). Indeed, Whittaker et al. (2005) have said that "It is disappointing that we still know so little about the power and timescale of 'species relaxation'." We agree. Our failure to answer this question makes predictions of species loss, as a consequence of habitat destruction and species invasions, very difficult to determine (Sax and Gaines 2008). Second, conflicting opinions about the future rate and speed of speciation as a consequence of habitat fragmentation and species invasions abound (e.g., Rosenzweig 2001; Vellend et al. 2007). Determining how likely we are to maintain the process of speciation and how it may be impaired or facilitated by human actions is of pressing concern, particularly in light of the many extinctions that we anticipate will occur in the future as a consequence of human actions. Ultimately, understanding both sides of the same coin, i.e., speciation and extinction, are fundamental to understanding ecological and evolutionary processes, but also fundamental to effectively conserving and promoting biological diversity. We believe that the continued study of ETIB, along with its embellishments, improvements, and

derivative theories are key to integrating ecological and evolutionary perspectives needed to best manage biological resources in the future.

Acknowledgements

This chapter benefited from detailed comments by Mark Lomolino, Rob Colwell and an anonymous reviewer. Consideration of ETIB in light of anthropogenic species invasions benefited from conversations with Jim Brown. Literature cited

- Allen, A. P., J. F. Gillooly, V. M. Savage, and J. H. Brown. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. Proceedings of the National Academy of Sciences, USA 103:9130-9135.
- Baker, H. G. and G. L. Stebbins. 1965. The genetics of colonizing species. Academic Press, New York.
- Blackburn, T. M., and K. J. Gaston. 2003. Macroecology: concepts and consequences: the 43rd Annual Symposium of the British Ecological Society. Blackwell Publishing, Malden.
- Brook, B. W., N. S. Sodhi, P. K. L. Ng. 2003. Catastrophic extinctions follow deforestation in Singapore. Nature 424:420-423.
- Brown, J. H. 1971. Mammals on mountaintops: nonequilibrium insular biogeography. American Naturalist 105:467-478.
- Brown, J. H. 1995. Macroecology. The University of Chicago Press, Chicago.
- Brown, J. H., and B. A. Maurer. 1989. Macroecology the division of food and space among species on continents. Science 243:1145-1150.
- Brown, J. H., S. K. M. Ernest, J. M. Parody, and J. P. Haskell. 2001. Regulation of diversity: maintenance of species richness in changing environments. Oecologia 126:321-332.
- Bruno, J. F., J. D. Fridley, K. D. Bromberg, and M. D. Bertness. 2005. Insights into biotic interactions from studies of species invasions. Pages 13-40 *in* D. F. Sax, S. D. Gaines, and J. J. Stachowicz, eds. Species invasions: Insights into ecology, evolution, and biogeography. Sinauer Associates, Sunderland.
- Cody, M. L. 2006. Plants on islands: diversity and dynamics on a continental archipelago. University of California Press, Berkeley.
- Currie, D. J. 1991. Energy and large-scale patterns of animal-species and plant-species richness. American Naturalist 137:27-49.
- Currie, D. J., and V. Paquin. 1987. Large-scale biogeographical patterns of species richness of trees. Nature 329:326-327.
- Dammermann, K. W. 1948. The fauna of Krakatau, 1883–1933. Koninklijke Nederlandsche Akademie Wetenschappen Verhandelingen 44:1–594.
- Darlington, P. J., Jr. 1957. Zoogeography: The geographical distribution of animals. John Wiley and Sons, New York.

Darwin, C. 1859. On the origin of species by means of natural selection. J. Murray, London.

- Diamond, J. M. 1969. Avifaunal equilibria and species turnover rates on the Channel Islands of California. Proceedings of the National Academy of Sciences, USA 64:57-63.
- Diamond, J. M. 1975. The island dilemma: lessons of modern biogeographic studies for the design of natural preserves. Biological Conservation 7:129-146.
- Eldredge, L.G., and S. E. Miller. 1995. How many species are there in Hawaii? Bishop Museum Occasional Papers 41:3–17.
- Enquist, B. J., J. H. Brown, and G. B. West. 1998. Allometric scaling of plant energetics and population density. Nature 395:163-165.
- Forster J. R. 1778. Observations Made during a Voyage Round the World, on Physical Geography, Natural History, and Ethic Philosophy. G. Robinson, London.
- Freckleton, R. P., and A. R. Watkinson. 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. Journal of Ecology 90:419-434.
- Funk, V. A., and W. L. Wagner. 1995. Hawaiian biogeography: evolution on a hot spot archipelago. Smithsonian Institution Press, Washington, DC.
- Gillespie, R. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science 303:356-359.
- Gottelli, N. J. 2004. Assembly Rules. Pages 1027-1121 in M. V. Lomolino, D. F. Sax, and J. H. Brown, eds. Foundations of biogeography: Classic papers with commentaries. The University of Chicago Press, Chicago.
- Grinnell, J. 1919. The English house sparrow has arrived in Death Valley: an experiment in nature. American Naturalist 53:468–472.
- Hanski, I. 1991. Single-species metapopulation dynamics concepts, models and observations. Biological Journal of the Linnean Society 42:14-38.
- Heaney, L. R. 2000. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. Global Ecology and Biogeography 9:59-74.
- Heaney, L. R. 2007. Is a new paradigm emerging for oceanic island biogeography? Journal of Biogeography 34:753-757.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.

- Huey, R. B., G. W. Gilchrist, M. L. Carlson, D. Berrigan, and L. Sierra. 2000. Rapid Evolution of a geographic cline in size in an introduced fly. Science 287:308-309.
- Kalmar, A., and D. J. Currie. 2006. A global model of island biogeography. Global Ecology and Biogeography 15:72-81.
- Kreft, H., W. Jetz, J. Mutke, G. Kier, and W. Barthlott. 2008. Global diversity of island floras from a macroecological perspective. Ecology Letters 11:116-127.
- Laurance, W. F. 2008. Theory meets reality: How habitat fragmentation research has transcended island biogeographic theory. Biological Conservation 141:1731-1744.
- Lawlor, T. E. 1986. Comparative biogeography of mammals on islands. Biological Journal of the Linnean Society 28:99-125.
- Lever, C. 1985. Naturalized mammals of the world. Longman, London.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15:237-240
- Levins, R. 1970. Extinction. Pages 75-108 *in* M. Gerstenhaber, ed. Some mathematical questions in biology. American Mathematical Society, Providence.
- Lomolino, M. V. 2000a. A call for a new paradigm of island biogeography. Global Ecology and Biogeography 9:1-6.
- Lomolino M. V. 2000b. A species-based theory of insular zoogeography. Global Ecology and Biogeography 9:39-58.
- Lomolino, M. V., and J. H. Brown. (in review) The reticulating phylogeny of island biogeography theory.
- Lomolino, M. V., J. H. Brown, and D. F. Sax. 2009. Island Biogeography Theory: Reticulations and re-integration of 'a biogeography of the species'. Pages xxx-xxx in J. B. Losos and R. E. Ricklefs, eds. The theory of island biogeography at 40: Impacts and prospects. Princeton University Press, Princeton.
- Losos, J. B., and R. E. Ricklefs. The Theory of Island Biogeography at 40: Impacts and Prospects. Princeton University Press, Princeton.
- Lynch, J. D., and N. V. Johnson. 1974. Turnover and equilibria in insular avifaunas, with special reference to the California Channel Islands. Condor 76:370-384.
- MacArthur, R. H., and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. Evolution 17:373-387.

- MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton.
- Mayr, E. 1965. Avifauna: turnover on islands. Science 150:1587-1588.
- O'Brien, E. M., R. Field, and R. J. Whittaker. Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. Oikos 89:588-600.
- Pressey, R. L., M. Cabeza, M. E. Watts, R. M. Cowling, and K. A. Wilson. 2007. Conservation planning in a changing world. Trends in Ecology and Evolution 22:583-592.
- Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.
- Rosenzweig, M. L. 2001. Loss of speciation rate will impoverish future diversity. Proceedings of the National Academy of Sciences, USA 98:5404-5410.
- Sax, D. F., B. P. Kinlan, and K. F. Smith. 2005. A conceptual framework for comparing species assemblages in native and exotic habitats. Oikos 108:457-464.
- Sax, D. F., J. J. Stachowicz, J. H. Brown, J. F. Bruno, M. N. Dawson, S. D. Gaines, R. K. Grosberg, A. Hastings, R. D. Holt, M. M. Mayfield, M. I. O'Connor, and W. R. Rice. 2007. Ecological and evolutionary insights from species invasions. Trends in Ecology and Evolution 22:465-471.
- Sax, D. F., and S. D. Gaines. 2003. Species diversity: From global decreases to local increases. Trends in Ecology and Evolution 18:561-566.
- Sax, D. F., and S. D. Gaines. 2006. The biogeography of naturalized species and the species-area relationship. Pages 449-480 in M. W. Cadotte, S. M. McMahon and T. Fukami, eds. Conceptual ecology and invasions biology: Reciprocal approaches to nature. Springer, Netherlands.
- Sax, D. F., and S. D. Gaines. 2008. Species invasions and extinction: the future of native biodiversity on islands. Proceedings of the National Academy of Sciences, USA 105:11490-11497.
- Sax, D. F., S. D. Gaines, and J. H. Brown. 2002. Species invasions exceed extinctions on islands worldwide: a comparative study of plants and birds. American Naturalist 160:766-783.
- Simberloff, D. 1974. Equilibrium theory of island biogeography and ecology. Annual Review of Ecology and Systematics 15:161-182.

- Simberloff, D., and E. O. Wilson. 1969. Experimental zoogeography of islands: The colonization of empty islands. Ecology 50:278-296.
- Simberloff, D., and E. O. Wilson. 1970. Experimental zoogeography of islands: A two year record of colonization. Ecology 51:934-937.
- Stachowitcz, J. J., and D. Tilman. 2005. Species invasions and the relationships between species diversity, community saturation, and ecosystem functioning. Pages 41-46 *in* D. F. Sax, S. D. Gaines and J. J. Stachowicz, eds. Species invasions: Insights into ecology, evolution, and biogeography. Sinauer Associates, Sunderland.
- Terborgh, J. 1973. Chance, habitat and dispersal in distribution of birds in West-Indies. Evolution 27:338-349.
- Terborgh, J. 1974. Preservation of natural diversity: The problem of extinction prone species. Bioscience 24:715-722.
- Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences, USA 101:10854-10861.
- Tilman D., R. M. May, C. L. Lehman, M. A. Nowak. 1994. Habitat destruction and the extinction debt. Nature 371:65-66.
- Toft, C. A., and T. W. Schoener. 1983. Abundance and diversity of orb spiders on 106 Bahamian islands: biogeography at an intermediate trophic level. Oikos 41:411-426.
- Turner, J. R. G., J. J. Lennon, and J. A. Lawrenson. 1988. British bird species distributions and the energy theory. Nature 335:539-541.
- Vellend, M., L. J. Harmon, J. L. Lockwood, M. M. Mayfield, A. R. Hughes, J. P. Wares and D. F. Sax. 2007. Effects of exotic species on evolutionary diversification. Trends in Ecology and Evolution 22:481-488.
- Whittaker, R. J. 1998. Island biogeography: Ecology, evolution, and conservation. Oxford University Press, Oxford.
- Whittaker, R. J. 2000. Scale, succession and complexity in island biogeography: Are we asking the right questions? Global Ecology and Biogeography 9:75-85.
- Whittaker, R. J. 2006. Island species-energy theory. Journal of Biogeography 33:11-12.
- Whittaker, R. J., and J. M. Fernandez-Palacios. 2007. Island biogeography: Ecology, evolution and conservation, 2nd Edition. Oxford University Press, New York.

- Whittaker, R. J., K. A. Triantis, and R. J. Ladle. 2008. A general dynamic theory of oceanic island biogeography. Journal of Biogeography 35:977-994.
- Whittaker, R. J., M. B. Araujo, P. Jelpson, R. J. Ladle, J. E. M. Watson, and K. J. Willis. 2005. Conservation biogeography: Assessment and prospect. Diversity and Distributions 11:3-23.
- Wilson, E. O. 1961. The nature of the taxon cycle in the Melanesian ant fauna. American Naturalist 95:169–193.
- Wilson, E. O. 1969. The theory of species equilibrium. Brookhaven Symposium in Biology 22: 38-47, a.k.a. Diversity and Stability in Ecological Systems, Brookhaven National Laboratory, Brookhaven.
- Wilson, E. O., and E. O. Willis. 1975. Applied biogeography. Belknap Press of Harvard University, Cambridge.
- Wright, D. H. 1983. Species-energy theory: And extension of species-area theory. Oikos 41:496-506.
- Wylie, J. L., and D. J. Currie. 1993. Species energy theory and patterns of species richness: I. Patterns of bird, angiosperm, and mammal richness on islands. Biological Conservation 63:137-144.

Table 10.1: Propositions of the Equilibrium Theory of Island Biogeography

Ecological Propositions:

- The rate of immigration of species (i.e. the arrival of new species) to an island decreases as the number of species that have arrived on an island increases; the rate reaches zero when all species that could colonize from an available pool of species have done so.
- 2. The rate of immigration of species to an island decreases with increasing isolation from a pool of potential colonists.
- 3. The rate of immigration of species increases with increasing island size.
- 4. The rate of extinction of species established on an island (of a given size) increases with increasing numbers of species.
- 5. The rate of extinction of species established on an island decreases with increasing size of an island.
- 6. The number of species on an island will be determined by an equilibrium between rates of immigration and extinction.
- 7. The rate of species turnover, i.e. change in species composition, will be determined by an equilibrium between rates of immigration and extinction.

Evolutionary Propositions:

- 8. In addition to immigration, the number of species on an island can be increased by speciation.
- 9. Speciation will only be important to an equilibrium in species number in a "radiation zone" found at the outward limits of species capacity for natural dispersal.
- 10. The distance to a "radiation zone" is taxon specific.
- 11. An equilibrium in species number reached ecologically is a "quasi-equilibrium" that can be increased over evolutionary time.

Figure 10.1. The Equilibrium Theory of Island Biogeography. **A**. The basic equilibrium model. An equilibrium number of species (S) is set by two opposing processes, immigration (I) and extinction (E). The rate of immigration decreases and the rate of extinction increases linearly with increasing richness; the rate of immigration reaches zero when the entire pool (P) of potentially immigrating species have arrived. **B**. The full equilibrium model. Rates of immigration and extinction are bowed downward to reflect various suppositions of natural history and population ecology (see text). Immigration rates on islands far (I_F) from a source are expected to be lower than those on near (I_N) islands. Extinction rates are expected to be higher on small (E_S) islands than on large (E_L) islands. In addition to different equilibrium numbers of species, differences in rates of species turnover (T) are expected to vary with the combination of immigration and extinction rates that characterize any given island.



Figure 10.2. The Equilibrium Theory of Island Biogeography in light of anthropogenic species invasions. Prehistoric immigration rates (I_1) are presumed to be much lower than the human increased rates of immigration (I₂) now observed on islands, in part because the total pool size has been increased from those species able to reach islands before human influence (P_1) versus the greatly expanded number of species currently in the pool (P_2) of potential immigrants. A. Extinction rates observed (E_0) for vascular plants and freshwater fishes have been extremely low to date following increases in immigration rates, such that the observed (and inferred) equilibrium number of species (S_0) is much higher than the historic equilibrium (S_1) . It is conceivable that the potential extinction rate (E_P) is actually much higher, but that long time-lags exist before extinctions are manifest, making it difficult to know the actual shape of the extinction curve or how many species will potentially exist at equilibrium (S_P) in the future. **B**. The number of land bird species currently occurring on oceanic islands (S_2) is similar to historic numbers (S_1) , in spite of a large turnover in species composition. ETIB can account for such a pattern if the increased immigration rate (I_2) has been matched by an increased extinction rate (E_2) . C. If land birds and non-volant mammals are considered as single combined group then the net increase in richness (S₂) observed on islands can be interpreted in light of an increased immigration rate (I_2) .



Chapter 11: Theory of Ecosystem Ecology Ingrid C. Burke and William K. Lauenroth

Where does the "ecosystem ecology" section of an ecology textbook belong - at the beginning, or the end? Ecosystem ecology is frequently viewed as the organizational scale that fits logically after individuals, populations, and communities, the "largest level" of ecology. For many, ecosystem ecology is the least appealing part of ecology because it includes more chemistry, physics, and math than many of the other subfields of ecology, few if any familiar or charismatic individual organisms, and an entire vocabulary that refers to the arcane details of nutrient cycles. We think this perception can as likely be credited to the way ecosystem ecology is presented as it can to its content. Our interests and enthusiasm for ecosystem ecology are directly traceable to the fact that it intellectually embraces all of the "levels" of organization in ecology, including evolutionary ecology, and that it addresses the energy and matter that is exchanged among organisms and their environment: it is the fabric that ties together all of ecology. Furthermore, and particularly relevant to the 21st century, ecosystem ecology is critical for understanding our relationships with the environment. The ecological importance of many current environmental problems, including disturbance, ecosystem ecology.

Reiners (1986) suggested that there are at least two models or theories of ecosystem ecology. One addresses ecosystem energetics, and is largely based on the second law of thermodynamics. The second deals with ecosystem stoichiometry, or how the fundamental ratios of elements in organisms control the distribution of elements in the environment. Reiners further speculated that a third theory might address the response of ecosystems to disturbances.

In this chapter, we seek to integrate ideas that relate to ecosystem energetics and stoichiometry, and to provide a framework for predicting the effects of disturbance on ecosystems. Our domain includes energy and matter cycling, at any scale, and our approach is grounded in systems theory. *We propose that the single most important process for understanding ecosystem structure and function is net ecosystem production.* Net ecosystem production (NEP) is the net result of carbon inputs and losses over a particular time period

 $NEP = NEE \pm F_{lateral} - (F_{disturbance} + F_{leach}).$

NEE, net ecosystem exchange, is the net of gross photosynthesis and ecosystem respiration, $F_{lateral}$ is the lateral movement of carbon into or out of the system, $F_{disturbance}$ is loss of carbon to such things are fire or harvest, and F_{leach} is leaching losses to ground water. NEE is the only component of NEP that is important across all ecosystem types and is the focus of our presentation. Ideas associated with net ecosystem production integrate key components of ecosystem ecology, including stoichiometry, energetics, and predicting and understanding the consequences of disturbance.

Below, we will introduce two theories that focus on energy and matter cycling, using net ecosystem production as the integrating function, with their constituent propositions. Each theory is linked to the general theory of ecology presented by Scheiner and Willig (Table 11.1). For each proposition, we present a brief synopsis of the historical importance of the idea, and a case study for the application of the idea to current ecological science.

NET ECOSYSTEM PRODUCTION, ENERGETICS, STOICHIOMETRY, ENVIRONMENT, AND EVOLUTION

Theory 1: The rates of net ecosystem production and its distribution in time and space both respond to and control patterns and levels of oxidation, temperature, ultraviolet radiation, and the distribution of elements, and ultimately have constrained the evolution of metabolic pathways.

Ecosystem ecology is firmly rooted in physics, chemistry, and biology. All ecological processes are constrained by the laws of thermodynamics. But because ecosystems are open with respect to energy through the process of photosynthesis, they store energy in carbon-carbon bonds against the forces of entropy. Nearly all of the energy available for organisms and their interactions is provided in these carbon-carbon bonds. Both autotrophs and heterotrophs rely upon these compounds as inputs to energy releasing metabolic cycles, and so the spatial and temporal pattern in the net production of reduced carbon compounds (net ecosystem production, or NEP, Figure 11.1) is the most fundamental attribute of ecosystems and of the biosphere, driving trophic dynamics, the cycling of all biologically active elements, and the distribution of key chemical and physical conditions including oxidation, temperature, and acidity, in time and space.

PATTERNS OF NET ECOSYSTEM PRODUCTION BOTH RESPOND TO AND DRIVE ELEMENT DISTRIBUTION AND ENVIRONMENT

Proposition 1a: The heterogeneous distribution of elements and chemical features of Earth and ecosystems (oxidation, temperature, acidity, and nutrient retention) causes there to be separation in time and space of organisms that are dominantly autotrophic and heterotrophic.

Proposition 1b: The separation in time and space of autotrophic and heterotrophic processes drives the distribution of biologically active elements and many key chemical features of Earth and ecosystems.

Patterns of NEP are complex, through time from seasons to millennia, and in space from individuals to the biosphere. The key components of NEP, photosynthesis and respiration, are each controlled to a different degree by the physical environment (water, temperature, light, oxygen, other nutrients, gravity, etc). For instance, total rates of heterotrophic respiration are most sensitive to oxygen availability (and other electron acceptors) and temperature. Photosynthesis, or net primary production, is most sensitive to light, water and nutrients (Figures 11.2a and 11.2b). Because physical conditions vary in time and space, the balance of heterotrophic respiration and autotrophic production do as well, resulting in variable accumulation of organic matter, as well as the consumption and production of oxygen and carbon dioxide.

As a consequence, during Earth's early history, as organic matter accumulated on the surface, an oxygenated atmosphere formed, stratospheric ozone layer produced, incident ultra violet radiation reduced, the greenhouse gas carbon dioxide was removed from the atmosphere, and global temperature dropped; in these changes are the roots of the "coevolution of climate and life" (Lovelock and Margulis 1974; Margulis and Lovelock 1974; Schneider and Londer 1984; Schopf 1983). The separation of autotrophic and heterotrophic processes continues to both respond to and drive element distribution and the key chemical features of Earth and ecosystems (oxidation, temperature, acidity, and nutrient retention). Areas with low water availability, low light, or low nutrient availability are limited in autotrophic potential, and all biological activity. Examples include deserts, at depth in lakes, and the open oceans respectively. Of all of the times and places where autotrophic and heterotrophic processes are differentially favored, perhaps most important are those in which autotrophic activity is high, but limitations in oxygen or (alternate electron acceptors) result in low rates of respiration, leading to accumulations of organic matter (positive NEP) such as occurred during the Paleozoic Era.

Application and advancement of these ideas at this time are most evident in studies of land-atmosphere interactions related to the ecosystem influences over global and regional climate change and variability. Net ecosystem production is strongly influenced by climate, by land use change (including changing patterns of disturbance, such as fire), and nitrogen deposition, and, as the theory states, these changes have strong feedbacks back to atmospheric dynamics and climate. Over the past 2 centuries, land use change and resulting negative net ecosystem production on a global scale has been responsible for an estimated 25-49% of the increase in atmospheric CO₂ concentration, with a decrease in relative impact today (closer to 12%; Brovkin et al. 2004), but still significant. Human use of fixed carbon compounds for food, fuel, and building materials, represents $\sim 24\%$ of potential net primary production (Haberl et al. 2007; Vitousek et al. 1986). The effects of future climate change on terrestrial ecosystem NEP under warming, or increased N deposition, are predicted to be negative decreasing NEP, but much is unknown (Field et al. 2007). The interactions of climate change and ecosystem dynamics will be partly dependent upon vegetation responses to what may be novel climates for species assemblages (Williams et al. 2007). Further, there are strong influences of terrestrial ecosystem structure and functioning on latent and sensible heat exchange, the hydrologic cycle, and resulting regional scale weather and climate, through biophysical feedbacks (e.g. (Bonan 2008)). These issues, related to land-atmosphere dynamics, represent key questions whose answers will advance not only ecosystem science, but also influence human responses to compelling global environmental issues.

Proposition 1c: Ecosystem function, through NEP, has constrained evolution. Rates of NEP and its distribution control rates of oxidation, global temperature, and the levels of ultraviolet radiation reaching Earth's surface, ultimately constraining the evolution of metabolic pathways.

Many authors have written about the co-evolution of climate and life (Lovelock and Margulis 1974; Schneider and Londer 1984; Schopf 1983) but few represent this as a consequence of ecosystem processes. Clearly, the metabolic pathways of the first organisms evolved under the constraints of an environment characterized by high ultraviolet radiation, low availability of oxidizing and reducing pairs, and few energy sources. The key feedback to altering the environment and changing the constraints over evolution of new metabolic pathways was positive global net ecosystem production. Since the beginning of life, many important alternate pathways for autotrophy and heterotrophy have evolved. In some ways, this proposition contradicts the above emphasis on asymmetric controls over autotrophy and heterotrophy: clearly, there are numerous pathways for heterotrophic respiration using electron acceptors other than oxygen, and these are a good deal older than aerobic respiration (Lurquin 2003). Early autotrophic pathways (chemo-autotrophs) and heterotrophic pathways (anaerobic) have both evolved under, and contributed to, this process. But the dominance of aerobic environments on Earth occurred because of the positive NEP since life formed. The associated buildup of oxygen, according to current theory, generated an ozone layer that reduced incident ultraviolet radiation and mutation rates, and led to the current biochemical dominance of aerobic respiration (Cockell and Blaustein 2001). Further, the positive NEP (since Earth formed) has altered global temperatures by lowering greenhouse gas concentrations.

There are numerous current applications and advances in ecosystem science related to the evolutionary significance of biochemical pathways. Many of the capabilities of microorganisms, for example, that developed during earlier environmental conditions, may be critically important for understanding and managing greenhouse gas concentrations from an ecosystem perspective. For instance, methane, an important greenhouse gas, is oxidized by microorganisms that likely evolved very early, when the atmosphere was predominantly reducing. Understanding the molecular genetics associated with this process could be important not only for understanding evolutionary ecology, but for understanding ecosystem functioning, and long term greenhouse gas concentrations (Hallam et al. 2004). Understanding the genetic structure of many other microbial processes may have great importance for managing ecosystem greenhouse gas fluxes (Zak et al. 2006).

Proposition 1d. [Stoichiometry] Because organisms have relatively constant stoichiometry (ratios of carbon to other elements), the pattern of carbon storage by ecosystems (net ecosystem production) provides the fundamental control over spatial and temporal patterns of element retention in ecosystems, element release from ecosystems, and the distribution of elements among ecosystems. Small differences in stoichiometry among organisms feed back to influence ecosystem function.

An important component to ecosystem theory since the 1950's has been that organisms have relatively constant element stoichiometry (or elemental composition), and that through their

activities, they control the element distribution within their environments (Elser et al. 2000; Redfield 1958; Reiners 1986).

We assert that the critical mechanism of this stoichiometric feedback between organisms and their environments occurs through NEP. The distribution of all biologically active elements can be explained by the fact that they are bonded to carbon skeletons and travel through the biota in concert with carbon: thus, the rate at which carbon skeletons are formed and broken (NEP) controls element cycling. Reiners (1986) described two major categories of organic compounds, protoplasmic and mechanical, which have distinct element ratios. For instance, protoplasmic compounds (enzymes, nucleic acids, etc) have high concentration of nitrogen, while mechanical or structural components (wood, bone, shell) have higher concentrations of materials that are resistant to decay (e.g. lignin), which tend to accumulate in soils and sediments. Thus, the type of organism and its tissue allocation influences element distribution. However, the key notion is that all elements, when biologically active, are bonded to carbon skeletons formed through autotrophy, and broken through heterotrophy. Thus, where and when NEP is positive, biologically active elements are stored, and when it is negative, they are released to inorganic forms, leaving them vulnerable to movement through gaseous or soluble form, or to being bound in mineral form (see below). To say this in its plainest possible form - elements cannot accumulate in ecosystems unless NEP is positive. As Aldo Leopold wrote in his essay "Odyssey",

"An atom at large in the biota is too free to know freedom; an atom back in the sea has forgotten it. For every atom lost to the sea, the prairie pulls another out of the decaying rocks. The only certain truth is that its creatures must suck hard, live fast, and die often, lest its losses exceed its gains." (Leopold 1949).

Over the past several decades, there have been many important advances related to the idea that disturbance to the spatial and temporal distribution of NEP alters fundamental environmental characteristics such as oxygen concentration, temperature, and element distribution and retention. Alternatively, disturbance to those environmental characteristics alters the spatial and temporal distribution of NEP. We offer three examples to elaborate these ideas that span local, regional, and global scales.

In the mid-1970's, forest ecologists and managers proposed that the most important predictor of nitrate in stream water was watershed NEP (Gorham et al. 1979; Vitousek and

Reiners 1975). This hypothesis has been tested and extended to other nutrient elements (phosphate) in forests as well as to stream, desert, and urban ecosystems (Gorham 1961; Lewis and Grimm 2007). Reductions in net ecosystem production allow inorganic forms of nutrients to be exported from ecosystems. By contrast, disturbances that increase NEP, such as water addition in dry areas or nitrogen deposition in areas with adequate water, promote storage of nutrient elements in organic form.

Ideas regarding stoichiometry have recently been important in advancing the understanding of terrestrial-aquatic linkages with respect to a key environmental problem: large-scale, culturally induced eutrophication of fresh waters and resulting hypoxia of coastal waters (Conley et al. 2009; Diaz and Rosenberg 2008; Schindler 2006). Eutrophication is well known to occur because increases in phosphorus and/or nitrogen disproportionately increase autotrophic activity, subsequently leading to times and places in which there is net heterotrophic activity, resulting in oxygen deficits (Howarth 2008). While reducing anthropogenic inputs of either element would seem to be a high priority for reducing the possibility of environmental collapse (Conley et al. 2009), recent results suggest that the balance between N inputs and P inputs is critical for predicting appropriate management strategy. (Schindler et al. 2008) illustrated using long-term, whole ecosystem lake experiments that reducing N inputs alone does not reduce eutrophication; cyanobacterial N fixation increases in response to reduced N, maintaining N:P ratios, and the eutrophic condition. (Carpenter 2008) and (Schindler and Hecky 2009) suggest that studies of N:P stoichiometry over large temporal and spatial scales, particularly in estuarine systems, are at the scientific forefront for understanding how to better mitigate hypoxia.

THEORY 2: DISTURBANCE AND ECOSYSTEMS

Among the most important areas for theory to contribute to applied problems today is in understanding and predicting the effects of disturbance on ecosystems. White and Pickett (1985) define disturbance as "...any relatively discrete event in time that disrupts ecosystem, community, or population structure and changes resources, substrate, or the physical environment." For the purposes of this paper, whose domain is energy and matter dynamics, we define disturbance as any alteration in net ecosystem production, including intentional land or water management, and unintentional management or changes in the natural disturbance cycle. Theory 2: The effects of disturbance on ecosystem matter and energy dynamics may be predicted by three fundamental characteristics: the above- and belowground distribution of biomass, the turnover time of the compartments, and the evolutionary history of the dominant organisms.

Ecosystem ecologists conceive of matter and energy storage and cycling using systems approaches (for example (Odum 1983). We conceptualize storage compartments, or pools, as somewhat arbitrary collections that are connected by flows of material or energy (fluxes). We say arbitrary because often our conceptualization of components or fluxes is not strictly measureable or even biologically accurate; for instance, the separation of live biomass from dead is not practically possible in soil systems, which are complex assemblages of root material, exudates, live and dead microbial biomass, micro and macro invertebrates, and detritus. However, the construction of simple (as well as complex) conceptual and simulation models of ecosystems has led to important theoretical advances and practical understanding.

We identify 4 general responses of ecosystems to perturbations: resistance, resilience, altered steady states, and instability (Aber and Melillo 2001; Holling 1973; Figure 11.3). Resilient ecosystems are characterized by being easily removed from their initial state but returning quickly. Systems characterized by resistance are difficult to remove from their original state, but once removed, recover very slowly. Altered steady states occur in ecosystems when irreversible changes occur but lead to reasonably "stable" new situations. Instability of ecosystems, or system components, may occur when changes are wrought that lead to the collapse of food webs (Post et al. 2002) or capability of systems to maintain key structural components (e.g. catastrophic soil loss). Avoiding irreversible changes requires understanding what types of process occur that may lead to each of these 4 responses.

Among the most useful concepts for quantifying and describing ecosystem function is turnover time (also called residence time), which is an estimate of the rate at which units of energy or matter pass through pools. Turnover time can be estimated as the size of the pool divided by either its rate of input, or its rate of input (often the pool is assumed to be at steadystate, so that the calculation is the same). Temporal dynamics of systems or their components are, as we will see below, closely linked to the turnover time of the compartments.

Proposition 2a: The resistance or resilience of a pool is determined in part by the turnover time. Pools with short turnover times tend to be resilient, that is, easily removed from

their initial state but quick to recover and those with long turnover times tend to be resistant – more difficult to perturb and recovering slowly from disturbance (Aber and Melillo 2001; Odum 1983)

Disturbances may decrease ecosystem storage by reducing gross NPP through killing live plant parts or whole plants (harvest, logging, fire), through export of live material (fuel or fodder), or through increasing ecosystem respiration (increase in temperature, draining to stimulate aerobic decomposition, etc.): all of these can result in negative net ecosystem production depending on the relative magnitudes of effects on autotrophy and heterotrophy. Ecosystems with a relatively long carbon turnover time (e.g., alpine tundra, arid ecosystems, boreal forest) will recover slowly, perhaps over hundreds of years. A rule of thumb is that recovery takes approximately 3 turnover times. Alternatively, ecosystems with rapid turnover times, such as tropical forests or annual grasslands, may recover more on the order of years to decades.

The theory is useful for predicting many other biogeochemical phenomena as well. Atmospheric pollutants, with turnover times on the order of days $(SO_2, 1.5 \text{ days})$, may be removed at a rate proportional to turnover time, such that if inputs were ceased acid deposition components would take only weeks or less to disappear, showing a high level of resilience. By contrast, tropospheric dimolecular oxygen (O₂, turnover time 4,500 years), characterized by a very large pool and very long turnover time, is highly resistant to change and is of essentially no concern for being substantialy influenced by human activities over several centuries. Interestingly, the same processes that have minimal impact on oxygen concentrations – fossil fuel combustion and land clearing - have major proportional impacts on the concentration of carbon dioxide (CO_2 , turnover time ~4-25 year; Wigley 2000) in the atmosphere. Pools with intermediate turn over times are of most concern, because they are responsive enough to be influenced by humans, yet resistant enough that they recover slowly. Most of the important greenhouse gases (N₂O, CH₄, CO₂, O₃) fall into this category of intermediate turnover, with recovery times substantially longer than the duration of any particular political administration. This traditional view of disturbance and maintainance of ecosystem function has become particularly relevant as ecologists have begun to focus on ecosystem services, or the importance of ecosystem processes such as carbon storage (net ecosystem production) to human welfare (Assessment 2005). With increasingly intensive natural resources extraction activities, for fossil fuels, forest products, water, and other resources, ecosystem restoration is critical to maintaining such ecosystem services (Hall and O'Connell 2007). Incorporating ecosystem theory into restoration and reclamation remains a substantial challenge for ecosystem ecology (Day et al. 2009).

Proposition 2b: The distribution of biomass (dead and/or alive) above and belowground is an important determinant of the resistance or resilience of an ecosystem to any particular disturbance.

Disturbances may be characterized by their degree of focus on aboveground vs. belowground biomass; for instance, fire, logging, and hurricanes focus on aboveground plant and detrital components. Ecosystems with a large proportion of stored biomass belowground, or with short turnover times, are resilient to the impacts of aboveground disturbances (Burke 2008). Fire or harvesting in perennial grasslands, resulting in removal of only aboveground biomass, tend to have minimal impacts on carbon storage or nutrient cycling. By contrast, those same disturbances in forested systems have major long-lasting impacts. Similarly, perturbations that include plowing, mining, urbanization, focusing on belowground or whole ecosystem components, have long term and practically irreversible impacts on carbon storage, for all ecosystems, even those with a large proportion of biomass stored belowground.

Proposition 2c: Ecosystems subjected to disturbances that fall within the evolutionary history of the dominant organisms are characterized by resistance or resilience to change; those subjected to disturbances outside the evolutionary history of the dominant organisms respond with either altered steady states or instability.

Ecosystem managers have realized that land use management strategies that are tuned to the evolutionary history of organisms may be sustainable, relative to those that are not. A single management type, for instance grazing by large herbivores, has different influences in ecosystems with a long evolutionary history of grazing (Serengeti, North American Great Plains) (Milchunas et al. 1988), compared to ecosystems with plants not adapted to heavy grazing (e.g., North American Great Basin) where livestock grazing may result in irreversible changes in species composition. Similarly, ecosystems whose fire history includes long fire intervals (e.g., *Pinus contorta* in the western U.S.) may be well adapted for fire suppression within time frames of several centuries. However, lower montane forested ecosystems of the Rocky Mountain western U.S., with a historical range of variability of fire frequency of 8-12 years (Schoennagel et al. 2004), may be dramatically altered by long-term fire suppression, the buildup of fuels and resulting catastrophic large fires.

This proposition is closely linked to recent work that evaluates the relationship between ecosystem function and biodiversity: what are the impacts of changes in species composition on ecosystem function (Tilman and Downing 1994; Tilman et al. 1997; Tilman et al. 1998)? Introductions of exotic species, or losses of keystone species, are additional examples of situations that represent changes outside the evolutionary history of ecosystems. Following introductions by invasive species, food web dynamics, dominant plant life form (D'Antonio and Vitousek 1992), fire dynamics and disturbance patterns, and the entire structure of ecosystems (Sala et al. 1996) can be changed to alternative steady states. Clearly, inclusion of evolutionary history and natural disturbance in our land management provides predictions: ecosystems subjected to disturbances that fall within the evolutionary history of the dominant organisms will be characterized by resistance or resilience, while those subjected to disturbances outside the evolutionary history of the dominant organisms will respond with either altered steady states or instability.

SUMMARY

Scientific disciplines with a rich history of theory (e.g. physics; Kragh 1999) are well recognized to have not only invigorated fast-growing fields with myriad hypotheses and advances, but have also had strong impacts on society. Ecosystem theory, embracing and integrating biology, evolutionary ecology, chemistry, physics, and global ecology, has had strong historical influence on natural resource management. At this critical juncture, when human societies are placing large demands on ecosystem services at local to global scales, ecosystem theory has much to offer for understanding ecosystem sustainably.

Literature cited

Aber, J., and J. Melillo. 2001, Terrestrial Ecosystems. New York, Harcourt Academic Press.

Assessment, M. 2005. Ecosystems and Well-Being: A synthesis, Island Press.

- Bonan, G. B. 2008. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. Science 320:1444-1449.
- Brovkin, V., S. Sitch, W. von Bloh, M. Claussen, E. Bauer, and W. Cramer. 2004. Role of land cover changes for atmospheric CO2 increase and climate change during the last 150 years. Global Change Biology 10:1253-1266.
- Burke, I. C., A. R. Mosier, P. B. Hook, D.G. Milchunas, J.E. Barrett, M.A. Vinton, R. L.
 McCulley, J. P. Kaye, R. A. Gill, H.E. Epstein, R. H. Kelly, W. J. Parton, C.M. Yonker, P.
 Lowe, and W.K. Lauenroth. 2008. Soil organic matter and nutrient dynamics of shortgrass steppe ecosystems *in* W. K. B. Lauenroth, I. C., ed. Ecology of the Shortgrass Steppe: A
 Long Term Perspective. New York, Oxford University Press.
- Carpenter, S. R. 2008. Phosphorus control is critical to mitigating eutrophication. Proceedings of the National Academy of Sciences of the United States of America 105:11039-11040.
- Cockell, C., and A. R. Blaustein. 2001, Ecosystems, evolution and ultraviolet radiation. New York Springer.
- Conley, D. J., J. Carstensen, R. Vaquer-Sunyer, and C. M. Duarte. 2009, Ecosystem thresholds with hypoxia:21-29.
- D'Antonio, C. M., and P. M. Vitousek. 1992. Biological invasions by invasive species, the grass fire cycle, and global change. Annual Review of Ecology and Systematics 23:63-87.
- Day, J. W., C. A. Hall, A. Yanez-Arancibia, D. Pimentel, C. I. Marti, and W. J. Mitsch. 2009. Ecology in Times of Scarcity. Bioscience 59:321-331.
- Diaz, R. J., and R. Rosenberg. 2008. Spreading dead zones and consequences for marine ecosystems. Science 321:926-929.
- Elser, J. J., R. W. Sterner, E. Gorokhova, W. F. Fagan, T. A. Markow, J. B. Cotner, J. F. Harrison et al. 2000. Biological stoichiometry from genes to ecosystems. Ecology Letters 3:540-550.
- Field, C. B., D. B. Lobell, H. A. Peters, and N. R. Chiariello. 2007. Feedbacks of terrestrial ecosystems to climate change. Annual Review of Environment and Resources 32:1-29.

- Gorham, E. 1961. Factors influencing supply of major ions to inland waters with special reference to the atmosphere. Geology Society of America Bulletin. 72:795-840.
- Gorham, E., P. Vitousek, and W. A. Reiners. 1979. The regulation of chemical budgets over the course of terrestrial ecosystem succession. Annual Review of Ecology and Systematics 10:53-84.
- Haberl, H., K. H. Erb, F. Krausmann, V. Gaube, A. Bondeau, C. Plutzar, S. Gingrich et al. 2007. Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. Proceedings of the National Academy of Sciences of the United States of America 104:12942-12945.
- Hall, J., and E. O'Connell. 2007. Earth systems engineering: turning vision into action.Proceedings of the Institution of Civil Engineers-Civil Engineering 160:114-122.
- Hallam, S. J., N. Putnam, C. M. Preston, J. C. Detter, D. Rokhsar, P. M. Richardson, and E. F. DeLong. 2004. Reverse methanogenesis: Testing the hypothesis with environmental genomics. Science 305:1457-1462.
- Holling, C. S. 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4:1-23.
- Howarth, R. W. 2008. Coastal nitrogen pollution: A review of sources and trends globally and regionally. Harmful Algae 8:14-20.
- Kragh, H. 1999, Quantum Generations: A history of physics in the twentieth century. Princeton NJ, Princeton University Press.
- Leopold, A. 1949, A Sand County Almanac. New York, Oxford University Press.
- Lewis, D. B., and N. B. Grimm. 2007. Hierarchical regulation of nitrogen export from urban catchments: Interactions of storms and landscapes. Ecological Applications 17:2347-2364.
- Lovelock, J. E., and L. Margulis. 1974. Atmospheric homeostasis by and for the biosphere: the Gaia hypothesis. Tellus 22:2-9.
- Lurquin, P. F. 2003, The origins of life and the universe. New York, Columbia University Press.
- Margulis, L., and J. E. Lovelock. 1974. Biological modulation of the Earth's atmosphere. Icarus 21:471-489.
- Milchunas, D. G., O. E. Sala, and W. K. Lauenroth. 1988. A generalized model of the effects of grazing by large herbivores on grassland community structure. The American Naturalist 132:87-106.

Odum, H. T. 1983. Systems ecology. New York, John Wiley & Sons, Inc.

- Post, J. R., M. Sullivan, S. Cox, N. P. Lester, C. J. Walters, E. A. Parkinson, A. J. Paul et al. 2002. Canada's recreational fisheries: The invisible collapse? Fisheries 27:6-17.
- Redfield, A. C. 1958. The biological control of chemical factors in the environment. The American Scientist Autumn:205-221.
- Reiners, W. A. 1986. Complementary models for ecosystems. The American Naturalist 127:59-73.
- Schindler, D. W. 2006. Recent advances in the understanding and management of eutrophication. Limnology and Oceanography 51:356-363.
- Schindler, D. W., and R. E. Hecky. 2009. Eutrophication: More Nitrogen Data Needed. Science 324:721-722.
- Schindler, D. W., R. E. Hecky, D. L. Findlay, M. P. Stainton, B. R. Parker, M. J. Paterson, K. G. Beaty et al. 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences of the United States of America 105:11254-11258.
- Schneider, S. H., and R. Londer. 1984, The coevolution of climate and life, Random House Books
- Schoennagel, T., T. T. Veblen, and W. H. Romme. 2004. The interaction of fire, fuels, and climate across rocky mountain forests. Bioscience 54:661-676.
- Schopf, J. W. 1983. Evolution of earth's earliest ecosystems. Recent progress and unsolved problems, Pages 361-364 *in* J. W. Schopf, ed. Earth's Earliest Biosphere. Princeton, NJ, Princeton University Press.
- Tilman, D., and J. A. Downing. 1994. Biodiversity and stability in grasslands. Nature 367:363-365.
- Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Siemann. 1997. The influence of functional diversity and composition on ecosystem processes. Science 277:1300-1302.
- Tilman, D., C. L. Lehman, and C. E. Bristow. 1998. Diversity-stability relationships statistical inevitability or ecological consequence. American Naturalist 151:277-282.
- Vitousek, P. M., P. R. Ehrlich, A. H. Ehrlich, and P. A. Matson. 1986. Human appropriation of the products of photosynthesis. Bioscience 36(6):368-373.

- Vitousek, P. M., and W. A. Reiners. 1975. Ecosystem succession and nutrient retention: a hypothesis. Bioscience. 25:376-381.
- White, P. S. and S. T. A. Pickett. 1985. Natural disturbance and patch dynamics: An introduction. Chapter 1 In: Pickett, S. T. A. and P. S. White (eds.) The ecology of natural disturbance and patch dynamics. Academic Press, Inc., Orlando.
- Wigley, T. M. L. S., D.S. 2000, The carbon cycle. Cambridge, Cambridge University Press.
- Williams, J. W., S. T. Jackson, and J. E. Kutzbacht. 2007. Projected distributions of novel and disappearing climates by 2100 AD. Proceedings of the National Academy of Sciences of the United States of America 104:5738-5742.
- Zak, D. R., C. B. Blackwood, and M. P. Waldrop. 2006. A molecular dawn for biogeochemistry. Trends in Ecology & Evolution 21:288-295.

General Ecology Fundamental	Theory 1: The rates of net ecosystem production and its
Principles, from Scheiner and	distribution in time and space both respond to and control
Willig And Laws of	patterns and levels of oxidation, temperature, ultraviolet
Thermodynamics	radiation, and the distribution of elements, and ultimately
	have constrained the evolution of metabolic pathways
Environmental conditions are	
heterogeneous in space and time.	[Patterns of net ecosystem production both
Deserves and finite and	environment]
heterogeneous in space and time	
neterogeneous in space and time.	and chemical features of Earth and ecosystems (oxidation,
Organisms are distributed in	temperature, acidity, and nutrient retention) causes there
space and time in a heterogeneous	to be separation in time and space of organisms that are
manner.	dominantly autotrophic and heterotrophic.
The distributions of according	
and their interactions depend on	Proposition 10: The separation in time and space of
contingencies	dutotrophic and heterotrophic organisms arives the
contingencies	chemical features of Earth and ecosystems
The ecological properties of	Proposition 1c: Ecosystem function through NEP has
species are the result of evolution	constrained evolution. Rates of NEP and its distribution
T	control rates of oxidation, global temperature, and the
	levels of ultraviolet radiation reaching Earth's surface,
	ultimately constraining the evolution of metabolic
	pathways
Organisms interact with their	Proposition 1d: [Stoichiometry] Because organisms have
abiotic and biotic environments.	relatively constant stoichiometry (ratios of carbon to other
	elements), the pattern of carbon storage by ecosystems (net
	ecosystem production) provides the fundamental control
	over spatial and temporal patterns of element retention in
	distribution of elements among ecosystems, and the
	differences in stoichiometry among organisms feed back to
	influence ecosystem function.
Laws of thermodynamics,	Proposition 1e: Each element has unique chemistry, that
conservation of matter.	contributes to its biological function, stoichiometric
	relationships, and distribution in the environment. Three
	aspects of elemental chemistry (below we refer to them as
	elemental observations) may be used to predict the
	differences among elements in their ecosystem cycling,
	their accumulation in ecosystems, and their impact on net
	ecosystem production.

Table 11.1. Relationship of Ecosystem Theory to General Ecology Principles.

General Ecology Fundamental Principles, from Scheiner and Willig and Laws of Thermodynamics	Propositions from Ecosystem Theory 2: Theory 2: The effects of disturbance on ecosystem matter and energy dynamics may be predicted by three fundamental characteristics: the turnover time of the compartments, the above- and belowground distribution of biomass and the evolutionary history of the dominant organisms.
Laws of thermodynamics, conservation of matter.	Proposition 2a: The resistance or resilience of a pool or a system is determined in part by the turnover time. Pools with short turnover times tend to be resilient, that is, easily removed from their initial state but quick to recover and those with long turnover times tend to be resistant – more difficult to perturb and recovering slowly from disturbance
Organisms interact with their abiotic and biotic environments.	Proposition 2b: The distribution of biomass (dead and/or alive) between above and belowground is an important determinant of the resistance or resilience of an ecosystem to any particular disturbance.
The ecological properties of species are the result of evolution	Proposition 2c: Ecosystems subjected to perturbations that fall within the evolutionary history of the dominant organisms are characterized by resistance or resilience to change; those subjected to perturbations outside the evolutionary history of the dominant organisms respond with either altered steady states or instability.



Figure 11.1. Net ecosystem production represents the net storage of energy as reduced carbon in ecosystems, and integrates of understanding of matter and energy in ecosystems.



Figures 11.2a and b. Net ecosystem production both responds to and controls element availability and environmental characteristics. Some conditions favor high net ecosystem production (high light, high nutrients, high water availability), and others negative net ecosystem production, by stimulating heterotrophy (high oxygen, high temperature). In return, high rates of net ecosystem production reduce oxygen availability, and high rates reduce nutrient availability.



Figure 11.3. There are 4 possible responses of ecosystems to perturbations: resilience, resistance, altered steady states, and instability. The responses may be predicted by the turnover time of the affected pools, and the evolutionary history of the organisms.

Chapter 12: Perspectives on Global Change Theory Debra P.C. Peters, Brandon T. Bestelmeyer, and Alan K. Knapp

Human influences on ecological drivers are increasingly recognized as dominant processes across a range of spatial and temporal scales. Regional to global-scale changes in drivers and important resources, such as atmospheric carbon dioxide concentrations, climate, and nitrogen deposition, are known to alter biotic structure, ecosystem function, and biogeochemical processes with feedbacks to human activities and the atmosphere (Petit et al. 1999; Grimm et al. 2000; Fenn et al. 2003; IPCC 2007). Human activities also directly affect ecosystems at finer scales through urbanization, species movement and extinction, and changes in land use that, in aggregate, have global impacts as human populations continue to increase and migrate (Alig et al. 2004; Theobold 2005; Grimm et al. 2008b). Although large volumes of data on global change drivers and ecological responses to them have been synthesized (e.g., SNE 2002; MEA 2005; Canadell et al. 2007; IPCC 2007), a coherent body of ecological theory focused on global change is lacking (Peters et al. 2008).

The U.S. Global Change Research Act of 1990 defined global change as: "Changes in the global environment -including alterations in climate, land productivity, oceans or other water resources, atmospheric chemistry, and ecological systems- that may alter the capacity of the Earth to sustain life". It is now clear that direct and indirect human actions are responsible for most of the more dramatic change occurring today and forecast for the future (Vitousek et al. 1997b). Global change is an aggregate of different forces that operate across all scales; many of these forces will be discussed in this chapter. In contrast, drivers historically studied by ecologists were assumed to occur locally and were ecosystem-specific, such as fire in forests and floods in streams. For contemporary ecological systems, a framework is needed to integrate across scales and to better understand the effects of multiple interacting drivers on ecosystem dynamics. For example, elevated concentrations of carbon dioxide (CO₂) and other greenhouse gases in the atmosphere are increasing temperatures globally whereas precipitation regimes are changing locally; ecosystem responses to these multiple, interacting drivers are often unknown (IPCC 2007). Existing ecological theories can be brought to bear on global change issues, but these theories need to be adapted and modified such that the key and sometimes unique aspects of global change drivers and responses to them are explicitly considered.
Ecosystem responses to global change drivers are often measured experimentally across a range of relatively fine scales, from individuals or portions thereof (e.g., leaves) to plots containing populations or communities (e.g., Shaw et al. 2002; Magill et al. 2004; Morgan et al. 2007; Siemann et al. 2007). These results are then extrapolated to broader scales (e.g., geographic distributions of species) using spatial patterns of responses or combined with phenological responses (Walther et al. 2002; Parmesan and Yohe 2003; Root et al. 2003). Alternatively, changes in regional- to global-scale patterns in vegetation structure and productivity have been estimated directly using simulation models or remotely sensed images (Defries et al. 2000; Sitch et al. 2008). In most cases, these studies have been conducted at one scale or in some cases, at multiple, independent scales (e.g., Peterson et al. 2003). Less attention has been devoted to how patterns and processes interact across scales to generate emergent behavior (e.g., Peters et al. 2004; Allen 2007). The propagation of fine-scale dynamics to larger scales of pattern often can not be predicted using linear extrapolation as a method for up-scaling. Similarly, the overwhelming effect of broad-scale drivers on fine-scale dynamics can not be understood by simply downscaling effects of these drivers (Peters et al. 2007). Alternative approaches are needed that account for cross-scale interactions (Peters et al. 2009). Because global changes in drivers and responses are inherently cross-scale and are connected spatially (Peters et al. 2008), theories of global change must account for these interactions.

There are many well-known examples of how fine-scale processes can propagate to influence large areas, and how broad-scale drivers can overwhelm fine-scale variation in pattern. For example, land use practices in central Asia, including overgrazing by livestock and cultivation of marginal lands, are interacting with effects of drought to result in high plant mortality and increased soil erosion at the scale of individually managed fields (e.g., 1-10 ha). Because most farmers and ranchers are following the same practices, these field-scale dynamics often aggregate nonlinearly with thresholds to the landscape and regional scales to generate large dust storms, the frequency of which has increased from 1 in 31 years to one per year starting in 1990 (Liu and Diamond 2005). As these dust storms continue to expand in spatial extent, they can travel inter-continentally to influence air quality in North America (<u>http://svs.gsfc.nasa.gov/goto?2957</u>), and can overwhelm natural determinants of local air quality (Jaffe et al. 2003).

In this chapter, we outline the characteristics of a theory of global change that draws upon a range of existing theories, including those from population biology and landscape ecology, as well as other disciplines such as Earth system sciences. We develop the basis for this theory and provide supporting evidence for its utility. We also provide an example where misleading results are likely to be obtained if the underlying concepts for such a theory are not accounted for, and we discuss new research directions based on this theory.

DOMAIN, PROPOSITIONS, AND MECHANISMS

The domain of a theory of global change is the causes and consequences of ecological properties of systems when the natural and human-induced drivers interact across a spatial and temporal hierarchy of scales. As a theory of global change develops, there are four key propositions that need to be considered - two occur under current conditions with natural variation in drivers, and two are unique to systems experiencing global change. These four propositions provide the context from which we derive our perspective on global change theory.

Our propositions follow from the eight fundamental principles of a theory of ecology, as proposed by Scheiner and Willig (2008, Chapter 1). In addition, spatial and temporal interactions of patterns and processes, originally articulated by Watt (1947), are a fundamental concept in our developing theory. Our perspective integrates a number of theories operating at specific scales with theories that link scales. We also incorporate ideas from other disciplines (physical sciences, human systems) to address connectivity by water, wind, and humans. We believe that any global change theory needs to be organized by spatial scales that correspond to scales through time.

Propositions under current conditions

In our developing theory, pattern-process relationships interact across a hierarchy of scales (Proposition 1) to result in spatial heterogeneity and connectivity among spatial units to be the critical system properties governing dynamics (Proposition 2). Our first proposition is derived from hierarchy theory (Allen and Star 1982) and a framework for interactions across scales (Peters et al. 2004, 2007). Our second proposition combines a framework developed for heterogeneous arid landscapes (Peters et al. 2006) with an emerging connectivity framework designed to relate fine-scale dynamics with broad-scale drivers (Peters et al. 2008).

Proposition 1. Interactions across scales. Our first proposition is that there is natural variation in environmental drivers and system responses that form a hierarchy of interacting spatial and temporal scales (Figure 12.1a). This proposition expands on general principles from hierarchy theory where a small number of structuring processes control ecosystem dynamics; each process operates at its own temporal and spatial scale (Allen and Starr 1982; O'Neill et al. 1986). Finer scales provide the mechanistic understanding for behavior at a particular scale, and broader scales provide the constraints or boundaries on that behavior. Other chapters in this book describe these fine-scale mechanisms more completely (Holt Chapter 7; Leibold Chapter 8) or describe the relationships among scales, such as the use of patches to explain landscape-scale dynamics in metapopulation theory (Hastings Chapter 6) and succession theory (Pickett et al. Chapter 9).

The concept of pattern-process interactions provides a general mechanism for dynamics within scales that lead to shifts in 'scale domains' (*sensu* Wiens 1989). Functional relationships between pattern and process are consistent within each domain of scale such that linear extrapolation is possible within a domain (Wiens 1989). Thresholds occur when pattern–process relationships change rapidly with a small or large change in a pattern or environmental driver (Groffman et al. 2006; Bestelmeyer 2006); both external stochastic events and internal dynamics can drive systems across thresholds (Scheffer et al. 2001).

Interactions among local and broader-scale processes can be important to patterns of distribution, abundance, and diversity (Ricklefs 1987; Levin 1992; Carpenter and Turner 2000). These cross-scale interactions generate emergent behavior that cannot be predicted based on observations at single or multiple, independent scales. For example, human activities at local scales can drive land change dynamics at regional scales (Luck et al. 2001; Dietz et al. 2007). Cross-scale interactions can also be important to metapopulation dynamics in that demographic and dispersal processes interacting with habitat heterogeneity can drive these dynamics across scales (Schooley and Branch 2007). A number of theories have used hierarchy theory as a basis for describing cross-scale interactions, including theories of complex systems (Milne 1998), self-organization (Rietkerk et al. 2004), panarchy (Gunderson and Holling 2002), and resilience (Holling 1992). Recently, a framework was developed to explain how patterns and processes at different scales interact to create nonlinear dynamics with thresholds (Peters et al. 2004, 2007). This framework focuses on the importance of connectivity and spatial heterogeneity in

determining how pattern-process relationships interact across scales, and forms the basis for our Proposition 2.

Proposition 2. Connectivity and spatial heterogeneity. Our second proposition is that intermediate-scale properties associated with connectivity and spatial heterogeneity determine how pattern-process relationships interact from fine to broad scales (Figure 12.1b). Within a domain of scale (i.e., fine, intermediate or broad), patterns and processes reinforce one another and are relatively stable. However, changes in drivers or disturbance can modify these patternprocess relationships in two ways: (1) Fine-scale patterns can result in positive feedbacks where new processes and feedbacks become important as the spatial extent increases. This change in dominant process is manifested as nonlinear, threshold changes in pattern and process rates. The nonlinear propagation of fire through time is one example where connectivity in fuel load shifts from individual tree morphology to within-patch distribution of overstory and understory plants to among-patch variability in topography and species distributions as the spatial extent of a fire expands (Allen 2007). (2) Broad-scale drivers can overwhelm fine-scale processes, such as regional drought that produces widespread erosion and minimizes the importance of local process such as competition to ecosystem dynamics. At the scale of landscapes, dispersal of invasive species can overwhelm local environmental variation in vegetation, soils, and grazing pressure to drive invasion dynamics (Peters et al. 2006).

Spatial heterogeneity and connectivity are inter-related: it is the combination of trends through time and patterns across space that lead to measures of connectivity. For example, expansion of an invasive species across a landscape can follow a nonlinear pattern where cover of the invasive increases through time for any given point within a spatial extent (Figure 12.1b[i]). At any given point in time, invasive species cover is heterogeneously distributed across the spatial extent in a number of different ways, from uniformly high or low or with a gradient or ecotonal spatial structure (Figure 12.1b[ii]). Combining trends through time with patterns in space lead to nonlinear changes in area dominated by the invasive species as a percentage of the spatial extent through time (Figure 12.1b[iii]). There are three thresholds in the system that are points in time where the slope of the line changes discontinuously as the dominant process changes (T1, T2, T3). A species that initially invades a landscape will first spread within a patch such that local recruitment of seeds and competition among plants are the dominant processes. As more seeds are produced and more plants succeed within a patch, a threshold is crossed where dispersal to other patches becomes increasingly important to landscape-scale pattern. The slope of each line segment (e.g., % invasive cover/ha/y) between each pair of thresholds (T1 – T2; T2 – T3) is a measure of the connectivity of plants of the invasive species across the landscape. The importance of spatial context is illustrated by the adjacency of each point to other points such that points closer to the area dominated by the invasive species are more likely to be invaded than points at farther distances.

Our connectivity proposition is itself based on the following propositions developed from Peters et al. (2007): (1) Global-scale patterns emerge from a hierarchy of interacting processes that propagate responses from fine to broad scales (i.e., plants to landscapes and regions). Finescale patterns often cannot be understood without knowledge of broader-scale processes. (2) Dynamics at any location on the globe are affected to varying degrees by transfer processes that connect adjacent as well as distant locations. (3) Transfer processes (wind, water, biota) connect locations via the movement of organisms, materials, disturbance, and information. The loss of historic transfer processes can result in disconnected locations. Conversely, an increase in magnitude and frequency of transfer processes can increase connectivity among previously isolated locations. (4) Spatial heterogeneity determines how drivers and transfer processes interact and feedback on one another across scales. (5) The relative importance of fine- or broadscale pattern-process relationships can vary through time, and alternate as the dominant factors controlling system dynamics.

We illustrate changes in cover through time and across space as estimates of connectivity and a description of the mechanisms producing these changes using a landscape change scenario from the northern Chihuahuan Desert in southern New Mexico, USA (Peters et al. 2004). A combination of field survey-based maps (1915, 1928/29), black-and-white aerial photos (1948), and pan-sharpened QuickBird satellite images (2003) were georegistered for a 942 ha pasture at the USDA ARS LTER site north of Las Cruces, NM (32°30'N, 106°48'W) (Figure 12.2a). Three classes of vegetation were digitized manually, and ARCGIS was used to obtain the area occupied by each of three classes through time: shrubs, grasses, and the ecotone between them.

In this landscape-scale example, most points on the landscape convert from grass-(Figure 12.2b) to shrub-dominated cover (Figure 12.2c) through time. At any point in time, in general, the pattern across the landscape changes from grass-dominated in the west-southwest (left side of panels in Figure 12.2a) to shrub-dominated in the east (right side of panels in Figure 12.2a). Aggregating this information to the entire landscape results in a nonlinear increase in area dominated by shrubs through time (Figure 12.2d). Three thresholds occur that are likely associated with a change in the dominant process driving dynamics across the landscape, from interspecific competition between individual plants in the early stages (prior to T1) to connections between shrubs by fine-scale water redistribution and long-distance seed dispersal. (T1-T2). Recruitment and seed dispersal (T2-T3) create connections among shrub patches as infilling occurs, although at a slower rate than the previous years. At later stages (T3 \rightarrow), the density and spatial arrangement of shrub patches result in low connectivity among isolated shrubs. In contrast, bare soil interspaces become highly connected by wind erosion to result in deposition of soil and nutrients under shrub canopies. These positive feedbacks to shrub persistence promote the development of dune fields that further limit success of grasses (Peters et al. 2004).

Additional propositions under global change

There are two additional propositions that are unique to and characterize phenomena considered under global change, and these in particular present unique challenges to existing theories (Figure 12.3b).

Proposition 3. Human drivers of global change. The third proposition is that human activities are ultimately the dominant drivers of global change (P3: Figure 12.3b top panel). The consequences of this proposition are that the dynamics and characteristics of key drivers previously recognized as governed by natural earth systems processes and feedbacks, such as atmospheric CO_2 and other greenhouse gases, climate, nitrogen deposition, are now determined to a large extent by human activities. These activities are a product of cultural, economic, and social systems (Pickett et al. 2001).

When combined with the widespread direct impacts of human actions on biological and ecological systems, this proposition means that many of the primary forces of change in ecology, as well as responses, interactions, and consequences of change, operate either partially or completely independent of evolutionary mechanisms, such as natural selection, that historically have been considered essential for understanding the ultimate basis of and context for ecological dynamics (Vitousek et al. 1997b; Palmer et al. 2004). For example, plant communities that exist in urban ecosystems do not necessarily reflect adaptations to the local environment – instead the

environment is often altered to permit the species to co-exist. Thus, the spatial patterns of individuals, their population dynamics, and overall productivity are largely a function of human activities and preference driven by socio-economic factors (Grimm et al. 2008a). Other examples include agricultural fields, water bodies devoted to aquaculture, planted "improved" pastures, and forest plantations in which the dominant species and their traits are no longer a product of evolutionary and ecological interactions, but instead are largely influenced by a human value system. Less obvious, but no less pervasive, is the attempted management and restoration (decidedly human activities) of natural areas to match environments and to achieve ecological states that may no longer exist (Hobbs et al. 2006), thus requiring significant resource inputs and human intervention (Seastedt et al 2008).

Proposition 4. Change in trajectories of global change drivers. Our fourth proposition is that global change drivers are of historically unprecedented magnitude, and as a consequence they are leading to trajectories of ecosystem responses that differ radically from those observed in the past (Figure 12.3b bottom panel). Increasing concentrations of atmospheric CO₂ and other greenhouse gases are primarily related to human activities, in particular fossil fuel emissions and land use change driven mainly by tropical deforestation (IPCC 2007). Atmospheric CO₂ concentrations have increased ca.100 ppm since 1750, and are currently higher than at least the past 650,000 years (Siegenthaler et al. 2005). These changes in atmospheric chemistry result in global temperature increases and regional increases or decreases in precipitation that often interact with changes in land cover to feed back to local weather (Pielke et al. 2002). Human activities also result in increases in nitrogen deposition in the form of nitrate from the combustion of fossil fuels and from ammonium, a by-product of animal metabolism and fertilization (Vitousek et al. 1997a; Fenn et al. 2003).

Recent studies predict that some future climates will have no historic analogs, and some extant climates may disappear (Fox 2007; Williams et al. 2007). These novel climates would likely result in new species associations (Hobbs et al. 2006), and the disappearance of climates could result in species extinctions (Overpeck et al. 1991). These "no analog" communities may result in ecological surprises with unknown responses to future climates (Williams and Jackson 2007). Additional global change drivers will likely interact with novel climates to result in even more surprising dynamics (Hobbs et al. 2006). One likely result of novel climates interacting with changes in other global drivers is that ecosystems will be pushed past critical thresholds to

result in irreversible ecosystem state changes. These state changes or regime shifts are increasingly recognized as important consequences of global change (Scheffer et al. 2001; Folke et al. 2004). Critical thresholds are often crossed either during or following a state change such that a return to the original state is difficult or seemingly impossible (Bestelmeyer 2006).

Combining these four propositions enables new predictions to be made about the effects of changing global drivers on ecosystem responses across scales. Because global change drivers have altered trajectories through time compared to historic dynamics, ecological responses through time at any point in space may have complex dynamics that may either continue to increase or even decrease (Figure 12.3c). For example, shrub cover in Figure 12.2 could continue to expand across arid and semiarid landscapes under conditions of increasing CO₂ concentrations and higher winter precipitation that favors shrub growth over grasses (Morgan et al. 2007). Alternatively, shrub cover could decrease if climatic changes favor grasses and increase fire frequencies (Briggs et al. 2005). Because the processes associated with shrub expansion (recruitment, competition, mortality) are not expected to change under global change, the general spatial pattern at any point in time is not expected to change (Figure 12.3d). Thus, one prediction is that changes in the temporal characteristics of global change drivers will impact arid landscapes more than changes in the spatial pattern of these drivers. The combination of altered temporal dynamics and no changes in spatial patterns can generate system responses with either increases, decreases or no changes in responses through time and space with thresholds that indicate changes in level of connectivity (Figure 12.3e). In addition, changes in spatial pattern may result via unexpected interactions in drivers and responses, such as the emergence of extreme climatic events, pest outbreaks, and altered disturbance regimes (Running 2008), that would result in even more complex or surprising behavior.

A framework for global change

The four propositions combine to form a conceptual framework that has connectivity as its foundation (Figure 12.4; modified from Peters et al. 2008). At the global scale, a hierarchy of interacting scales governs dynamics. Dynamics at any one location on the globe depend on both local patterns and processes at that location, and the movement of materials via transfer processes from other locations. All places on Earth are connected through a globally mixed atmosphere and regionally through a variety of biotic and abiotic mechanisms, such as human

transport of propagules, toxins, and diseases as well as propagation of disturbances and changes in land use as influenced by global economics. Thus, changes in one location can have dramatic influences on both adjacent and non-adjacent areas, either at finer or broader scales. Transfer processes associated with the movement of air, water, animals, and humans provide these connections, both within and across scales (Figure 12.2).

Disruptions in connectivity are also becoming increasingly important, often in different parts of a system that are increasing in connectivity. For example, land use practices over the past several centuries have increased the density of corn and soybean fields in the Midwestern U.S. to result in a highly connected mosaic of agricultural fields. In contrast, the plowing of tallgrass prairie for agricultural fields has resulted in disconnected remnant prairie locations throughout the region. As a result, movement of agricultural pests and disease among fields is facilitated, but the movement of plants and animals between fragmented remnant prairies is difficult because of the large distances between fragments.

Transfer processes and spatial heterogeneity can either amplify or attenuate system response to broad-scale drivers. Amplification occurs when the rate of change in system properties increases. This increase can result from high spatial heterogeneity or homogeneity that promotes cascading events, such as the nonlinear spread of wildfires (Peters et al. 2004). Cascading events in which a fine-scale process propagates nonlinearly to have an extensive impact have also been documented in the climate system, in lakes, and in the invasion of perennial grasslands by woody plants (Lorenz 1964; Peters et al. 2004; Wilson and Hrabik 2006). Attenuation occurs when the rate of change decreases through time, such as the decrease in wave amplitude as the wave form associated with a tsunami increases (Merrifield et al. 2005). The result is that the greatest effects of a tsunami occur closest to the source of the seismic event, and spatial heterogeneity in land or sea features become increasingly important as distance from the seismic event increases (Fernando and McCulley 2005). The spread of wildfires also attenuates with time and with decreases in fuel load or changes in weather conditions. In addition, broad-scale drivers, such as drought, can act to overwhelm fine-scale variation in vegetation, topography, and soils to result in homogeneous responses over large areas (Albertson and Weaver 1942).

The relative importance of fine- or broad-scale pattern-process relationships can vary through time, and compete as the dominant factors controlling system dynamics. For example,

connectivity of larvae from coral reef fishes is more locally important and regionally more variable than previously thought based on new analyses of dispersal constraints interacting with physical oceanography (Cowen et al. 2006). Processes that connect spatial units, such as dispersal of woody plants, are important under some conditions whereas local processes, such as soil texture, dominate on other sites; in both cases, broad-scale drought can overwhelm finer-scale processes to result in similar dynamics during dry years (Yao et al. 2006).

INSIGHTS FROM CONNECTIVITY PERSPECTIVE

We illustrate the importance of a connectivity-based theory of global change for addressing one specific ecological problem: the effects of hurricanes on ecological systems. We first show how our first two propositions apply to current conditions for both drivers and ecosystem responses. We then show how the two global change specific propositions (anthropogenic origin of drivers and changes in trajectories of drivers) are needed to understand and predict the impacts of hurricanes within the context of other global change drivers in the future.

Current conditions

Drivers of hurricane activity. Although it is readily accepted that hurricanes are disturbances with major impacts on ecosystems, the drivers controlling the formation, intensity, and track of hurricanes remain poorly understood. Recent research indicates that hurricane development is affected by drivers and processes operating across a range of spatial and temporal scales (Proposition 1) and that these drivers and processes interact such that spatial heterogeneity and connectivity among spatial units are important (Proposition 2). Because these propositions are related, we discuss them together within the context of hurricane development.

Physical processes interacting within and among scales predominate in the development of hurricanes (Goldenberg et al. 2001); these processes are directly or indirectly affected by global change drivers. Hurricanes that affect North America most often start as small thunderstorms in the Western Sahel region of Africa and increase in spatial extent and intensity as they propagate westward across the Atlantic Ocean (Dunn 1940; Landsea 1993). Both local factors, such as sea surface temperatures within the region of hurricane development, and broadscale factors, such as the El Niño-Southern Oscillation (ENSO) in the tropical Pacific and continental precipitation in West Africa determine whether or not an African thunderstorm develops into a hurricane (Gray 1990; Glantz et al. 1991; Landsea and Gray 1992; Saunders et al. 2000; Donnelly and Woodruff 2007).

Ecological responses to hurricanes. Hurricanes impact ecosystems across a range of spatial and temporal scales that influence ecosystem responses. Spatial pattern of damage resulting from hurricanes is scale-dependent: at the scale of individual trees and small stands, tree age and height, species composition, stand structure and soil conditions influence amount and type of damage (Foster and Boose 1992; Ostertag et al. 2005). At broader scales, spatial variability in vegetation, land use, environmental conditions, and disturbance history are important as well as landscape- and watershed-scale factors that connect spatial units, such as wind speed and direction, precipitation intensity, and topographic gradients (Boose et al. 2001; Sherman et al. 2001). Recent studies suggest that ecosystem properties, such as stand age and condition, forest type, aspect, and landscape-scale measures of connectivity, including distance to nearest perennial stream, are more important predictors of forest damage patterns than broad-scale drivers of wind speed and duration (Kupfer et al. 2008).

Ecosystem responses following hurricanes are also scale-dependent (reviewed in Everham and Brokaw 1996; Lugo 2008), and can include interactions across scales as a result of changes in connectivity among spatial units (Willig et al. 2007). For example, landscape reconfiguration and disruption of dispersal among patches by hurricanes can interact with local demographics of species to influence patterns in biodiversity across scales (Willig et al. 2007).

Global change conditions

Drivers of hurricanes, although incompletely understood, are being influenced by anthropogenic sources of variation (Proposition 3) and the trajectories of these drivers are changing (Proposition 4). We focus on both local and broad-scale drivers that are likely to change.

Local drivers. Sea surface temperatures [SSTs] have increased nonlinearly over the 20th century in the Atlantic Ocean (Trenberth 2005). This trend has been attributed to global warming and human activities (IPCC 2007). In addition, the amount of total column water vapor over the global oceans has increased 1.3% per decade (Trenberth 2005). Both higher SSTs and increased water vapor tend to increase energy available for atmospheric convection and thunderstorm production that can lead to hurricane development. There is general agreement that human-induced environmental changes occurring in hurricane regions can increase hurricane intensity

and rainfall (Emanual 2005; Goldenberg et al. 2001; Webster et al. 2005). There is less agreement on predicted effects of global warming on hurricane frequency with unclear evidence that frequencies are changing beyond the range of historic variation (Henderson-Sellers et al. 1998; Goldenberg et al. 2001).

Broad-scale drivers. Over the past 5000 years, the frequency of intense hurricane landfalls on centennial to millennial time scales was likely related to variations in ENSO and the strength of the West African monsoon (Donnelly and Woodruff 2007). Thus, reliable forecasts of remote conditions will be needed for predicting the occurrence and intensity of hurricanes in North America (Pielke and Landsea 1999). In addition, nonlinearities in the climate system that lead to threshold dynamics may make predictions based on historic trends difficult and unreliable (Rial et al. 2004).

For ENSO, a 155 year reconstruction from the tropical Pacific shows a gradual transition in the early 20th century and an abrupt shift in 1976 to new periodicities that reflected changes in the regional climate towards warmer and wetter conditions (Urban et al. 2000). The dramatic shift in 1976 coincided with a global shift in temperatures attributed to anthropogenic global warming (Graham 1995; Mann et al. 1998). Thus, global warming could further alter the frequency of ENSO cycles. However, additional factors need to be considered, such as sharp decreases in sea surface temperatures in the Atlantic Ocean interacting with dust-induced feedback processes that can moderate hurricane intensity (Lau and Kim 2007).

Factors that influence future rainfall in the Western Sahel will undoubtedly affect the number, duration, and intensity of hurricanes in North America (Webster et al. 2005). Climate projections for this region are unclear: one model predicts severe drying in the latter part of the 21st century while another predicts wet conditions throughout this time period, and a third model predicts modest drying (Cook and Vizy 2006). Clearly, better climate predictions and an understanding of the relationship between rainfall and wave formation are needed before rainfall on the continent of Africa can be used to predict hurricanes in the North Atlantic.

New insights to ecological systems

Atmospheric scientists have known since at least the 1960s that hurricanes connect the African and North American continents (Gray 1968). However, the complex cross-scale interactions determining hurricane development, intensity, and movement track have only recently been recognized and better appreciated as critical elements of a connected Earth system. A sense of urgency in understanding and prediction now predominates in the literature, in particular as our global environment continues to change and as hurricane damage increases with population density and wealth along U.S. coastlines (Pielke and Landsea 1999).

Ecological systems will continue to be influenced by hurricanes, both in obvious and subtle ways because of connections that link spatial and temporal scales (Hopkinson et al. 2008). Ecosystems in the track of hurricanes along the coast of North America are composed of species that evolved under the current hurricane regime, and it remains to be seen how different parts of these systems and different ecosystem types will respond as hurricane activity changes in the future (Michener et al. 1997). Even ecosystems located outside the direct path of hurricanes can be affected by a change in disturbance regime: deserts in southern New Mexico received within several days ca. 43% of the annual average rainfall as a result of the remnants of Hurricane Dolly in 2008. These extreme, remote events are not included in climate change projections for these systems (Seager et al. 2007), yet an increase in hurricane activity would reverse the direction of these projections from drier to wetter. Because these deserts have undergone dramatic changes from grasslands to shrublands over the past 150 years that are at least partially related to drought cycles, an increase in rainfall may provide opportunities for some landscape locations to revert to grass dominance, a state change that is considered unlikely under current climatic conditions.

Furthermore, hurricanes are not the only disturbance with local and global drivers that impact ecological systems across scales (Dale et al. 2001). Hurricanes are often associated with other disturbance events that accentuate the effects of wind and rain: drought often follows hurricanes with its own effects on surviving organisms (Covich et al. 2006). Fire can follow hurricanes with greater impacts on birds than the hurricane itself (Lynch 1991).

RESEARCH DIRECTIONS

Future research should include a consideration of drivers and responses interacting across multiple spatial and temporal scales, and explicitly measure transport processes when they are important to dynamics. Five characteristics of systems have been identified to account for multiple scales of variation (reviewed in Peters et al. 2006): (1) local processes (e.g., recruitment, competition, and mortality) interacting with microsite environmental variability (e.g., climate, soils, disturbance history), (2) historic legacies that influence the local environmental conditions,

the current assemblage of species and their ability to respond, (3) current environmental drivers and (4) future environmental drivers with local to global-scale variability, and (5) transport processes that connect spatial units across a range of scales, from the landscape to the globe.

It is the spatial scaling of characteristics 3-5 that ecologists need to consider when studying ecological problems within the context of global change, yet this aspect has received the least amount of attention to date. A consideration of variability in drivers from remote locations, such as rainfall in West Africa that influences hurricane activity in North America, is often missing from ecological studies. Although there is increasing recognition of the importance of ENSO and other climatic cycles on local rainfall patterns, the interaction of these cycles with other drivers (e.g., elevated CO₂, nitrogen deposition) has not typically been considered, although climatic cycles are related to disturbances, such as wildfire (Kitzberger et al. 2007).

Because transport processes are typically time and labor intensive to measure, it is important that an initial step in ecological studies is to determine where, when, and how transport processes *may be* important relative to the other drivers in order to decide if sampling is justified. An important aspect of this developing global change theory and its associated framework is that transport processes need to be considered within the context of the properties of the system and the questions to be addressed, but they do not need to be explicitly sampled for all questions.

Both direct and indirect drivers, such as disturbances, have characteristic spatial and temporal scales that need to be studied as interactive effects on ecosystems. Wildfires, floods, insect outbreaks, and other episodic events are not yet included in climate change models (Running 2008), yet they often interact across scales to result in surprising ecosystem responses (e.g., Allen 2007; Ludwig et al. 2007; Young et al. 2007).

Recently, new approaches to studying continental-scale problems under global change have been presented (e.g., Crowl et al. 2008; Grimm et al. 2008b; Hopkinson et al. 2008; Marshall et al. 2008; Williamson et al. 2008). Here we summarize three key recommendations from these papers that are relevant across scales. <u>First</u>, existing long-term datasets can be used to compare trends across sites. Similar patterns in data through time for sites located throughout a region or in different parts of the continent can suggest the presence of a global driver determining synchronicity in these dynamics, such as observed with wildfires and climatic cycles (Kitzberger et al. 2007). Alternatively, similar internal processes may be controlling system dynamics in different locations to overwhelm variation in drivers. A major limitation to these multi-site analyses has been accessibility of comparable data. Recent efforts at synthesizing long-term data and metadata from many U.S. sites are allowing these comparisons to be conducted (e.g., <u>http://www.ecotrends.info</u>). <u>Second</u>, coordinated efforts are needed to explicitly examine the importance of fine- to broad-scale drivers and transport processes to ecosystem dynamics across many sites. Existing networks of sites need to be coordinated such that comparable data are collected and compared dynamically in order to identify connections among sites, and to predict effects of cascading events as they influence adjacent and non-contiguous areas, such as the impacts of wildfires on air and water quality in burned sites and at distant locations. <u>Third</u>, simulation models are needed to complement experiments in order to provide more complete spatial and temporal coverage of ecosystem responses to global change drivers. Process-based models will be required to forecast a future with conditions that are unprecedented in Earth's history. In contrast, an empirical extrapolation of responses based on current or past conditions will result in large uncertainty. Multi-disciplinary approaches will be needed to account for the complexity of interactions across scales and levels of organization in the ecological hierarchy.

SUMMARY

Direct and indirect drivers of ecological systems are changing nonlinearly in response to human activities. These drivers and ecological systems interact across a range of spatial and temporal scales that often result in non-intuitive ecosystem dynamics. We outline some basic propositions to frame the development of a theory of global change based on connectivity within and among adjacent and non-contiguous spatial units. This nascent theory builds on several more mature bodies of theory developed for specific scales or levels of organization, and uses the concept of cross-scale interactions based on transport processes to link several of these theories. As our knowledge of the interacting components of the Earth System expands with improvements in sensing, measuring, and modeling technologies, we expect corresponding refinements to the theory that will improve its coherence and utility.

ACKNOWLEDGMENTS

This study was supported by National Science Foundation awards to LTER programs at the Jornada Basin at New Mexico State University (DEB 0618210) and the Konza Prairie at Kansas State University (DEB 0218210).

Literature Cited

- Albertson, F.W., and J.E. Weaver. 1942. History of the native vegetation of western Kansas during seven years of continuous drought. Ecological Monographs 12: 23-51.
- Alig, R.J., J.D. Kline, and M. Lichtenstein. 2004. Urbanization on the U.S. landscape: looking ahead in the 21st century. Landscape and Urban Planning 69:219-234.
- Allen, C.D. 2007. Interactions across spatial scales among forest dieback, fire, and erosion in northern New Mexico landscapes. Ecosystems 10: 797-808.
- Allen, T.F.H., and T.B. Starr. 1982. Hierarchy theory: perspectives for ecological complexity. University of Chicago Press, Chicago, IL, USA.
- Bestelmeyer, B.T. 2006. Threshold concepts and their use in rangeland management and restoration: the good, the bad, and the insidious. Restoration Ecology 14: 325-329.
- Boose, E.R., K.E. Chamberlin, and D.R. Foster. 2001. Landscape and regional impacts of hurricanes in New England. Ecological Monographs 71: 27-48.
- Briggs, J.M., A.K. Knapp, J.M. Blair, J.L. Heisler, G.A. Hoch, M.S. Lett, J.K. McCarron. 2005. An ecosystem in transition: causes and consequences of the conversion of mesic grassland to shrubland. BioScience 55: 243-254.
- Canadell, J.G., D.E. Pataki, and L.F. Pitelka, eds. 2007. Terrestrial ecosystems in a changing world. Springer, New York.
- Carpenter, S.R., and M.G. Turner. 2000. Hares and tortoises: Interactions of fast and slow variables in ecosystems. Ecosystems 3: 495-497.
- Cook, K.H., and E.K. Vizy. 2006. Coupled model simulations of the West African monsoon system: 20th and 21st century simulations. Journal of Climate 19: 3681-3703.
- Covich, A.P., T.R. Crowl, and T. Heartsill Scalley. 2006. Effects of drought and hurricane disturbances on headwater distributions of palaemonid river shrimp (Macrobrachium spp.) in the Luquillo Mountains, Puerto Rico. Journal North American Benthological Society 25: 99-107.
- Cowen, R.K., C.B. Paris, and A. Srinivasan. 2006. Scaling of connectivity in marine populations. Science 311: 522-527.
- Crowl, T.A., T.O. Crist, R.R. Parmenter, G. Belovsky, and A.E. Lugo. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. Frontiers in Ecology and the Environment 6: 238-246.

- Dale, V.H., L.A. Joyce, S. McNulty, R.P. Neilson, M.P. Ayres, M.D. Flannigan, P.J. Hanson,
 L.C. Irland, A.E. Lugo, C.J. Peterson, D. Simberloff, F.J. Swanson, B.J. Stocks, and B.M.
 Wotton. 2001. Climate change and forest disturbances. BioScience 51: 723-734.
- Defries, R.S., M.C. Hansen, J.R.G. Townshend, A.C. Janetos, and T.R. Loveland. 2000. A new global 1-km dataset of percentage tree cover derived from remote sensing. Global Change Biology 6: 247-254.
- Dietz, T., E.A. Rosa, and R. York. 2007. Driving the human ecological footprint. Frontiers Ecology Environment 5:13-18.
- Donnelly, J.P., and J.D. Woodruff. 2007. Intense hurricane activity over the past 5,000 years controlled by El Niño and the West African monsoon. Nature 447: 465-468.
- Dunn, G.E. 1940. Cyclogenesis in the tropical Atlantic. Bulletin of the American Meteorological Society 21: 215-229.
- Emanuel, K. A. 2005: Increasing destructiveness of tropical cyclones over the past 30 years. Nature 436: 686-688.
- Everham, E.M. III, and N.V.L. Brokaw. 1996. Forest damage and recovery from catastrophic wind. Botanical Review 62: 113-185.
- Fenn, M.E., R. Haeuber, G.S. Tonnesen, J.S. Baron, S. Grossman-Clarke, D. Hope, D.A. Jaffe,
 S. Copeland, L. Geiser, H.M. Rueth, and J.O. Sickman. 2003. Nitrogen emissions,
 deposition, and monitoring in the western United States. BioScience 53:391-403.
- Fernando, H.J.S., and J.L. McCulley. 2005. Coral poaching worsens tsunami destruction in Sri Lanka. Eoc 86(33):301-304.
- Folke, C., S.R. Carpenter, B. Walker, M. Scheffer, T. Elmquist, L. Gunderson, and C.S. Holling. 2004. Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review Ecology Systematics 35:557-581.
- Foster, D.R., and E.R. Boose. 1992. Patterns of forest damage resulting from catastrophic wind in central New England, USA. Journal of Ecology 80: 79-98.
- Fox, D. 2007. Back to the no-analog future? Science 316:823-825.
- Glantz, M.H., R.W. Katz, and N. Nicholls. 1991. Teleconnections linking worldwide climate anomalies: scientific basis and societal impact. Cambridge University Press, Cambridge. 534pp.

- Goldenberg, S.B., C.W. Landsea, A.M. Mestas-Nuñez, and W.M. Gray. 2001. The recent increase in Atlantic hurricane activity: causes and implications. Science 293: 474-479.
- Graham, N.E. 1995. Simulation of recent global temperature trends. Science 267: 686-691.
- Gray, W.M. 1968. Global view of the origin of tropical disturbances and storms. Monthly Weather Review 96:669-700.
- Gray, W.M. 1990. Strong association between West African rainfall and U.S. landfall of intense hurricanes. Science 249: 1251-1256.
- Grimm, N.B., Grove, J.M., Pickett, S.T.A. and C.L. Redman. 2000. Integrated approaches to long-term studies of urban ecological systems. Bioscience 50: 571–584.
- Grimm, N.B., S.H. Faeth, N.E. Golubiewski, C.L. Redman, J. Wu, X. Bai, and J.M. Briggs. 2008a. Global change and the ecology of cities. Science 319: 756-760.
- Grimm, N.B., D. Foster, P. Groffman, J.M. Grove, C.S. Hopkinson, K. Nadelhoffer, D.E. Pataki, and D.P.C. Peters. 2008b. The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. Frontiers in Ecology and the Environment 5: 264-272.
- Groffman, P.M., Baron, J.S., Blett, T., Gold, A.J., Goodman, I., Gunderson, L.H., et al. 2006. Ecological thresholds: the key to successful management or an important concept with no practical application? Ecosystems 9:1-13.
- Gunderson, L.H., and C. Holling, eds. 2002. Panarchy: understanding transformations in human and natural systems. Island Press, Washington, DC.
- Hastings, A. Theories of population dynamics. Pp in Scheiner, S.M., and M.R. Willig, eds. Theories in ecology.
- Henderson-Sellers, A., H. Zhang, K. Emanuel, W. Gray, C. Landsea, G. Holland, J. Lighthill, S-L. Shieh, P. Wabster, and K. McGuffie. 1998. Tropical cyclones and global climate change: a post-IPCC assessment. Bulletin of the American Meteorological Society 79: 19-38.
- Hobbs, R.J., S. Arico, J. Aronson, J.S. Baron, P. Bridgewater, V.A. Cramer, P.R. Epstein, J.J.
 Ewel, C.A. Klink, A.E. Lugo, D. Norton, D. Ojima, D.M. Richardson, E.W. Sanderson, F.
 Valladares, M. Vila, R. Zamora, and M. Zovel. 2006. Novel ecosystems: theoretical and management aspects of the new ecological world order. Global Ecology and Biogeography 15: 1-7.

- Holling CS 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4: 1-23.
- Holling, C.S. 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecological Monographs 62: 447-452.
- Holt, R. Enemy-victim interactions. Pp in Scheiner, S.M., and M.R. Willig, eds. Theories in ecology.
- Hopkinson, C.S., A.E. Lugo, M. Alber, A.P. Covich, and S.J. Van Bloem. 2008. Forecasting effects of sea-level rise and windstorms on coastal and inland ecosystems. Frontiers in Ecology and the Environment 6: 255-263.
- IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.
- Jaffe, D., I. McKendry, T. Anderson, and H. Price. 2003. Six "new" episodes of trans-Pacific transport of air pollutants. Atmospheric Environment 37: 391-404.
- Kitzberger, T., P.M. Brown, E.K. Heyerdahl, T.W. Swetnam, and T.J. Veblen. 2007. Contingent Pacific-Atlantic Ocean influence on multicentury wildfire synchrony over western North America. Proceedings of the National Academy of Sciences 104: 543-548.
- Kupfer, J.A., A. T. Myers, S.E. McLane, and G.N. Melton. 2008. Patterns of forest damage in a southern Mississippi landscape caused by Hurricane Katrina. Ecosystems 11: 45-60.
- Landsea, C.W. 1993. A climatology of intense (or major) Atlantic hurricanes. Monthly Weather Review 121: 1703-1713.
- Landsea, C.W., and W.M. Gray. 1992. The strong association between Western Sahel monsoon rainfall and intense Atlantic hurricanes. Journal of Climate 5: 435-453.
- Lau, K.M., and J.M. Kim. 2007. How nature foiled the 2006 hurricane forecasts. Eos Transactions, AGU: 88: 105-107.
- Leibold, M. Metacommunity theory. Pp in Scheiner, S.M., and M.R. Willig, eds. Theories in ecology.
- Levin, S.A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-67.

- Liu, J. and J. Diamond. 2005. China's Environment in a Globalizing World How China and the Rest of the World Affect Each Other. Nature 435:1179-1186.
- Lorenz, E.N. 1964. The problem of deducing the climate from the governing equations. Tellus XVI:1-11.
- Luck, M.A., G.D. Jenerette, J. Wu, and N.B. Grimm. 2001. The urban funnel model and spatially heterogeneous ecological footprint. Ecosystems 4:782-796.
- Ludwig, J.A., R. Bartley, A.A. Hawdon, B.N. Abbott, and D. McJannet. 2007. Patch configuration non-linearly affects sediment loss across scales in a grazed catchment in northeast Australia. Ecosystems 10: 839-845.
- Lugo, A.E. 2008. Visible and invisible effects of hurricanes on forest ecosystems: an international review. Austral Ecology 33: 368-398.
- Lynch, J.F. 1991. Effects of Hurricane Gilbert on birds in a dry tropical forest in the Yucatan Peninsula. Biotropica 28: 577-584.
- Magill, A.H., J.D. Aber, W.S. Currie, et al. 2004. Ecosystem response to 15 years of chronic nitrogen additions at the Harvard Forest LTER, Massachusetts, USA. Forest Ecology, and Management 196:7-28.
- Mann, M.E., R.S. Bradley, and M.K. Hughes. 1998. Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392: 779-787.
- Marshall, J.D., J.M. Blair, D.P.C. Peters, G. Okin, A. Rango, and M. Williams. 2008. Predicting and understanding ecosystem responses to climate change at continental scales. Frontiers in Ecology and the Environment 6: 273-280.
- Merrifield MA, Firing YL, Aarup T, Agricole W, Brundit G, Chang-Seng D, Farre R, Kilonsky B, Knight W, Kong L, Magori C, Manurung P, McCreery C, Mitchell W, Pillary S, Scindele F, Shillington F, Testut L, Wijeratne EMS, Caldwell P, Jardin J, Nakahara S, Porter F-Y, Turetsky N. 2005. Tide gauge observations of the Indian Ocean tsunami, December 26, 2004. Geophys Res Lett 32:L09603, doi:10.1029/2005GL022610.
- Michener, W.K., E.R. Blood, K.L. Bildstein, M.M. Brinson, and L.R. Gardner. 1997. Climate change, hurricanes and tropical storms, and rising sea level in coastal wetlands. Ecological Applications 7: 770-801.
- Millenium Ecosystem Assessment (MEA). 2005. Ecosystems and human well-being: synthesis. Island Press, Washington, DC.

- Milne, B.T. 1998. Motivation and benefits of complex systems approaches in ecology. Ecosystems 1:449-456.
- Morgan, J.A., D.G. Milchunas, D.R. LeCain, M. West, and A.R. Mosier. 2007. Carbon dioxide enrichment alters plant community structure and accelerates shrub growth in the shortgrass steppe. Proceedings National Academy Sciences 104:14724-14729.
- O'Neill, R.V., D.L. DeAngelis, J.B. Waide, and T.F.H. Allen. 1986. A hierarchical concept of ecosystems. Princeton University Press, Princeton, NJ, USA.
- Ostertag, R., F.N. W.L. Silver, and A.E. Lugo. 2005. Factors affecting mortality and resistance to damage following hurricanes in a rehabilitated subtropical moist forest. Biotropica 37: 16-24.
- Overpeck, J.T., P.J. Bartlein, and T. Webb III. 1991. Potential magnitude of future vegetation change in eastern North America: comparisons with the past. Science 254:692-695.
- Palmer, M., E. Bernhardt, E. Chonesky, S.L. Collins, A. Dobson, C. Duke, B. Gold, R. Jacobson, S. Kingsland, R. Kranz, M. Mappin, M.L. Martinez, F. Micheli, J. Morse, M. Pace, M. Pascual, S. Palumbi, O.J. Reichman, A. Townsend, and M. Turner. 2004. Ecology for a crowded planet. Science 304:1251-1252.
- Parmesan, C, and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37-42.
- Peters, D.P.C., B.T. Bestelmeyer, A.K. Knapp, J.E. Herrick, H.C. Monger, and K.M. Havstad. 2009. Approaches to predicting broad-scale regime shifts using changing pattern-process relationships across scales. Pp 47-72 in Miao, S., S. Carstenn, and M. Nungesser, eds. Real world ecology: large-scale and long-term case studies and methods. Springer, New York.
- Peters, D.P.C., B.T. Bestelmeyer, J.E. Herrick, H.C. Monger, E. Fredrickson, and K.M. Havstad. 2006. Disentangling complex landscapes: new insights to forecasting arid and semiarid system dynamics. BioScience 56:491-501.
- Peters, D.P.C., B.T. Bestelmeyer, and M. G. Turner. 2007. Cross-Scale Interactions and Changing Pattern-Process Relationships: Consequences for System Dynamics. Ecosystems 10:790-796.
- Peters, D.P.C., P.M. Groffman, K.J. Nadelhoffer, N.B. Grimm, S.L. Collins, W.K. Michener, and M.A. Huston. 2008. Living in a connected world: a framework for continental-scale environmental science. Frontiers in Ecology and the Environment 5: 229-237.

- Peters, D. P. C., R. A. Pielke, Sr, B. T. Bestelmeyer, C. D. Allen, S. Munson-McGee, and K. M. Havstad. 2004. Cross scale interactions, nonlinearities, and forecasting catastrophic events. Proceedings National Academy Sciences 101:15130-15135.
- Peterson, G. D., S. R. Carpenter, and W. A. Brock. 2003. Uncertainty and management of multistate ecosystems: an apparently rational route to collapse. Ecology 84:1403–1411.
- Petit, J.R., J. Jouzel, D. Raynaud, N.I. Barkov, J.M. Barnola. 1999. Climate and atmospheric history of the past 420000 years from the Vostok ice core, Antarctica. Nature 399:429-436.
- Pickett, S.T.A., M.L. Cadenassso, J.M. Grove, C.H. Nilon, R.V. Pouyat, W.C. Zipperer, and R. Costanza. 2001. Urban ecological systems: linking terrestrial ecological, physical and socioeconomic components of metropolitan areas. Annual review of Ecology and Systematics 32: 127-157.
- Pickett, S.T.A. Succession theory. Pp in Scheiner, S.M., and M.R. Willig, eds. Theories in ecology.
- Pielke R.A. Sr., G. Marland, R.A. Betts, T.N. Chase, J.L. Eastman, J.O. Niles, D.D.S. Niyogi and S.W. Running. 2002. The influence of land-use change and landscape dynamics on the climate system – relevance to climate change policy beyond the radiative effect of greenhouse gases. Philosophical Transactions of Royal Society of London Ser. A 360:1705-1719.
- Pielke, R.A., Jr., and C.N. Landsea. 1999. La Niña, El Niño, and Atlantic hurricane damages in the United States. Bulletin of the American Meteorological Society 80: 2027-2033.
- Rial, J.A., R.A. Pielke Sr, M. Beniston, M. Claussen, J. Canadell, P. Cox, H. Held, N. de Noblet-Ducoudé, R. Prinn, J.F. Reynolds, and J.D. Salas. 2004. Nonlinearities, feedbacks, and critical thresholds within the Earth's climate system. Climatic Change 65: 11-38.
- Ricklefs, R.E. 1987. Community diversity: relative roles of local and regional processes. Science 235: 167-171.
- Rietkerk, M, Dekker S.C., de Ruiter, P.C., and J. van de Koppel. 2004. Self-organized patchiness and catastrophic shifts in ecosystems. Science 305: 1926-1929.
- Root, T.L., J.T. Price, K.R. Hall et al. 2003. Fingerprints of global warming on wild animals and plants. Nature 421:57-60.
- Running, S.W. Ecosystem disturbance, carbon, and climate. Science 321: 652-653.

- Saunders, M.A., R.E. Chandler, C.J. Merchant, and F.P. Roberts. 2000. Atlantic hurricanes and NW Pacific typhoons: ENSO spatial impacts on occurrence and landfall. Geophysical Research Letters 27: 1147-1150.
- Scheffer, M., S.R. Carpenter, J.A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature 413:591-596.
- Scheiner, S.M., and M.R. Willig. 2008. A general theory of ecology. Theoretical Ecology 1:21-28.
- Scheiner, S.M., and M.R. Willig. A general theory of ecology. Pp in Scheiner, S.M., and M.R. Willig, eds. Theories in ecology.
- Schooley, R.L., and L.C. Branch. 2007. Spatial heterogeneity in habitat quality and cross-scale interactions in metapopulations. Ecosystems 10: 846-853.
- Seager, R., M. Ting, I. Held, Y. Kushnir, J. Lu, G. Vecchi, H-P Huang, N. Harnik, A. Leetmaa, N-C Lau, C. Li, J. Velez, and N. Naik. 2007. Model projections on an imminent transition to a more arid climate in southwestern North America. Science 316:1181-1184.
- Seastedt, T.R., R.J. Hobbs, and K.N. Suding. 2008. Management of novel ecosystems: are novel approaches required? Frontiers in Ecology and the Environment 6, doi:10.1890/070046.
- Shaw, M.R., E.S. Zavaleta, N.R. Chiariello, E.E. Cleland, H.A. Mooney, and C.B. Field. 2002. Grassland responses to global environmental changes suppressed by elevated CO₂. Science 298:1987-1990.
- Sherman, R.E., T.J. Fahey, and P. Martinez. 2001. Hurricane impacts on a mangrove forest in the Dominican Republic: damage patterns and early recovery. Biotropica 33: 393-408.
- Siegenthaler, U., T.F. Stocker, E. Monnin, D. Lüthi, J. Schwander, B. Stauffer, D. Raynaud, J.-M. Barnola, H. Fischer, V. Masson-Delmotte, J. Jouzel. 2005. Stable carbon cycle-climate relationship during the late Pleistocene. Science 310:1313-1317.
- Siemann, E., W.E. Rogers, and J.B. Grace. 2007. Effects of nutrient loading and extreme rainfall events on coastal tallgrass prairies: invasion intensity, vegetation responses, and carbon and nitrogen distribution. Global Change Biology 13:2184-2192.
- Sitch, S., C. Huntingford, N. Gedney, P.E. Levy, M. Lomas, S.L. Piao, R. Betts, P. Ciais, P. Cox,
 P. Friedlingstein, C.D. Jones, I.C. Prentice, and F.I. Woodward. 2008. Evaluation of the
 terrestrial carbon cycle, future plant geography, and climate-carbon cycle feedbacks using
 five Dynamic Global Vegetation models (DGVMs). Global Change Biology 14: 2015-2039.

The State of the Nation's Ecosystems (SNE). 2002. Measuring the lands, waters, and living resources of the United States. The H.J. Heinz III Center for Science, Economics, and the Environment. Cambridge University Press, New York.

Theobold, D.M. 2005. Landscape patterns of exurban growth in the USA from 1980 to 2020. Ecology and Society 10:32. [online] URL:http://www.ecologyandsociety.org/vol10/iss1/art32/.

- Trenberth, K.E. 2005. Uncertainty in hurricanes and global warming. Science 308: 1753-1754.
- Urban, F.E., J.E. Cole, and J.T. Overpack. 2000. Influence of mean climate change on climate variability from a 155-year tropical Pacific coral record. Nature 407: 989-993.
- Vitousek, P.M., J.D. Aber, R.W. Howarth, G.E. Likens, P.A. Matson, D.W. Schindler, W.H. Schlesinger, and D.G. Tilman. 1997a. Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications 7:737-750.
- Vitousek, P.M., H.A. Mooney, J. Lubchenco, and J.M. Melillo. 1997b. Human domination of Earth's ecosystems. Science 277:494-499.
- Walther, G-R, E. Post, P. Convey et al. 2002. Ecological responses to recent climate change. Nature 416:389-395.
- Watt, A. S. 1947. Pattern and process in the plant community. Journal of Ecology 35: 1-22.
- Webster, P.J., G.J. Holland, J.A. Curry, and H.R. Chang. 2005. Changes in tropical cyclone number, duration, and intensity in a warming environment. Science 309: 1844-1846.
- Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385-397.
- Williams, J.W., S.T. Jackson, and J.E. Kutzbach. 2007. Projected distributions of novel and disappearing climates by 2100 AD. Proceedings National Academy of Sciences 104:5738-5742.
- Williams, J.W., and S.T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Frontiers in Ecology and the Environment 9:475-482.
- Williamson, C.E., W. Dodds, T.K. Kratz, and M.A. Palmer. 2008. Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. Frontiers in Ecology and the Environment 6: 247-254.
- Willig, M.R., C.P. Bloch, N. Brokaw, C. Higgens, J. Thompson, and C.R. Zimmerman. 2007. Cross-scale responses of biodiversity to hurricane and anthropogenic disturbance in a tropical forest. Ecosystems 10: 824-838.

- Wilson, KA and TR Hrabik. 2006. Ecological change and exotic invaders. pp 151-167 in JJ Magnuson, TK Kratz and BJ Benson (eds), Long-Term Dynamics of Lakes on the Landscape. Oxford University Press, Oxford, England.
- Yao, J., D.P.C. Peters, K.M. Havstad, R.P. Gibbens, and J.E. Herrick. 2006. Multi-scale factors and long-term responses of Chihuahuan Desert grasses to drought. Landscape Ecology 21: 1217-1231
- Young, D.R., J.H. Porter, C.M. Bachmann, G. Shao, R.A. Fusina, J.H. Bowles, D. Korwan, and T.F. Donato. 2007. Cross-scale patterns in shrub thicket dynamics in the Virginia Barrier Complex. Ecosystems 10: 854-863.

Figure captions

Figure 12.1. Two propositions upon which a global change theory is based are occurring under current conditions: natural variation in system properties (drivers and responses) and connectivity among spatial units that link dynamics through time and across space. Connectivity is represented by the slope of each line segment between each pair of thresholds (T1-T2, T2-T3). Species cover is used as an example; other response variables are possible.

Figure 12.2. An example of a measure of connectivity in shrubs in the northern Chihuahuan Desert. Area covered by one of three vegetation types: shrubs, grasses, and the ecotone between them, was calculated for four dates using vegetation types and aerial photos on the left, and displayed through time in the graph for shrubs only. Three points in time and space were found where the rate of change in area increased nonlinearly to indicate a threshold (T1, T2, T3). The slope of each line segment between thresholds is a measure of the connectivity of shrubs over that time period. Insets show homogeneous plant cover [green in (b)] when the area is dominated by grasses. Under shrub dominance, patches of shrubs [green in (c)] are disconnected by bare interspaces that allow erosion by wind and water. Adapted from Peters et al. (2004).

Figure 12.3. Our emerging theory of global change includes the two propositions under current conditions (P1. variation in system properties, P2. connectivity) combined with P3. global change drivers that are experiencing P4. changing trajectories through time. The result is that connections among spatial units are increasing in both magnitude and frequency of occurrence primarily as a result of changes to trajectories through time rather than substantive changes in pattern.

Figure 12.4. Connectivity framework. Global change drivers and natural drivers are influencing a hierarchy of scales of system properties and responses. Each scale has its characteristic patterns and processes, and scales interact to generate changing pattern-process relationships through time and across space. Downscaling occurs when broad-scale drivers overwhelm fine-scale pattern (e.g., hurricanes), and upscaling occurs when fine-scale patterns and processes propagate to influence large spatial extents (e.g., dust plumes from field-based landuse practices. Redrawn from Peters et al. (2008).



Figure 12.1.











Figure 12.4

Chapter 13: A Theory of Ecological Gradients Gordon Fox, Samuel M. Scheiner and Michael R. Willig

Understanding the heterogeneous nature of species distributions is central to ecology as embodied in the first fundamental principle of its general theory (Table 1.3). As early as the 18th century, it was noted that species richness differed across the globe (von Humboldt 1808; Hawkins 2001). Today, it is widely recognized that species richness changes along a variety of gradients. Some gradients are spatial (e.g., latitude, depth, elevation; Willig et al. 2003), but may reflect underlying or correlated environmental variation (e.g., solar insolation with latitude). In this chapter, we do not discuss gradients from a purely spatial perspective, leaving such consideration to Colwell (Chapter 14). Here, we focus on gradients of species richness that pertain to environmental characteristics (e.g., disturbance, salinity, moisture; Grace 1999). Gradients with respect to productivity are probably the most widely discussed of these ecological gradients (Waide et al. 1999). As productivity increases, species richness may increase, decrease, assume a hump-shape or evince a U-shape, and the pattern may change with geographical or ecological scale (Mittelbach et al. 2001; Gillman and Wright 2006).

Our goal is to further develop a constitutive theory of environmental gradients of species richness as first promulgated by Scheiner and Willig (2005). We expand on that effort by refining the propositions of that theory, revealing hitherto concealed assumptions, and providing a conceptual framework that further unifies seemingly disparate models. We also examine an oft-cited model in detail, show that it is interpreted incorrectly by many, and discuss approaches for revising it. Besides improved understanding of the particular theory examined here, this exercise illustrates the process of theory development, emphasizing its dynamic nature.

DOMAIN OF THE THEORY AND ITS MODELS

The domain of our constitutive theory is environmental gradients in species richness. Literally, *gradient* refers to the slope of a curve; in this case the curve is richness as a function of some environmental characteristic. Slopes range from zero to infinity. Although most thinking about these gradients has concerned continuous variation in the environment (so that richness describes a smooth curve with continuous derivatives), there is no logical, biological, or mathematical reason why this must be so. Indeed, one can imagine a limiting case: a threshold in some environmental variable *x*, such that locations with $x < x_{crit}$ have dramatically lower richness than locations with $x > x_{crit}$. Rather than a smooth curve, the graph will be flat except at x_{crit} , where the change in richness will be immediate and represented by a vertical line. Mathematically, this is described by a step function, which has a slope of 0 everywhere but at the step itself, where the slope is infinite. While real ecological examples are probably less extreme than this, the step function is instructive in that it makes clear that the theory of species richness gradients can include very sudden changes in richness in ecological space, as well as gradual changes in richness between continuously varying environments. Models used to study richness under continuous environmental variation are likely to take a different form than those used for a small number of discretely different environments; we focus on the continuous case unless otherwise stated.

Ecological gradients occur in spatial contexts, but the theory itself is not necessarily spatial; in its broadest sense, the theory refers to species richness as a function solely of some environmental characteristic. Most intuitively think of ecological gradients as occurring over space, like the gradient from drier to wetter soils that occurs along a hillside. This intuition can be misleading, as the theory encompasses environmental variation occurring in any spatial or temporal pattern, on any spatial or temporal scale. The gradient need not be spatially contiguous or arranged so that the most similar environments are nearest to each other. For example, the theory may apply to a landscape consisting of randomly distributed patches in which environmental characteristics do not show spatial auto-correlation. The pattern of spatial or temporal contiguity and auto-correlation, or lack thereof, can determine which models are appropriate for consideration in any particular situation. The models that we consider here are not spatially explicit, although spatially-explicit versions are possible.

That said, particular models may be relevant only to particular spatial or temporal scales in that it is likely that different mechanisms (e.g., competition, speciation) will dominate at particular scales. The appropriate scale of a model generally depends on the assumptions of the model itself and on the biology of the taxa under consideration, rather than on an a priori scale. A critical distinction is whether the set of sites or collections of species under consideration draws on organisms from a single pool of species (a metacommunity; Leibold Chapter 8), or from multiple pools. Again, the importance of various ecological processes will differ in these instances.

Although the constitutive theory that we describe can apply to gradients at any scale, our focus in this chapter will be models with domains at the regional scale (10s to 100s of km^2) and with mechanisms operating in ecological time. These are the spatio-temporal scales for which most models have been developed concerning environmental gradients in species richness. To clarify the relationship between the theory of ecological gradients and spatial issues, it is useful to consider the relationship between gradient theory and species-area theory. Both theories involve predictions of richness as a function of another variable: resource or stressor concentration in gradient theory, and area in species-area theory (Figure 13.1). Richness is, of course, a function of both area and resources or stressors; however, we cannot yet draw a surface connecting the two-dimensional graphs in Figure 13.1 unless we assume that there is no interaction between area and resources or stressors. Currently both of these theories only permit limited views of such a relationship: gradient theory predicts richness as a function of resources or stressors for a fixed area (along a single plane slicing the three-dimensional space perpendicular to the area axis). Changing assumptions about the landscape (how patches of different resource levels are arranged in space, relative to dispersal processes) lead to different models under gradient theory. By contrast, species-area theory predicts richness along a single plane slicing the three-dimensional space perpendicular to the resource axis, and also requires assumptions about landscape-level variation. This suggests that a complete theory of species richness may have landscape-related dimensions in addition to the area and resource or stressor dimensions. The identities and number of these axes represents a problem that has yet to be explored.

Although a number of models examine aspects of richness gradients, few have clearly defined the relevant characteristics of the species pool. Models often fail to indicate whether taxonomic or ecological attributes delimit the species under consideration. For example, analyses may focus on all species within a clade and a particular level in the taxonomic hierarchy (e.g., a family), a functional guild (e.g., diurnal foliage-gleaning insectivores), an ensemble (e.g., frugivorous bats), or a trophic level (e.g., herbivores). There has been little consideration as to when a model should concern one or the other of these species pools. A species pool may also have a spatial component that is often ill-defined (e.g., biome specific or continental). A related issue is whether there is a single pool or a number of different pools, which again may be strictly

a spatial phenomenon (e.g., a gradient that stretches over multiple continents) or be ecological (e.g., a gradient that involves clades that specialize on different conditions).

As we show later, extant models treat species as identical in resource requirements, dispersal ability, and extinction probability. Clearly this violates something that is probably better documented than any other fact in ecology–species differ from one another. In practice, the models make this assumption, but in the literature, the species are assumed only to be roughly equivalent. Most authors refer to these models – or to related species-abundance models – as applying to limited groups such as particular taxa or guilds. For example, herbaceous annual plants might be thought of as roughly equivalent, whereas herbaceous plants and trees are certainly not, because herbs and trees have very different mortality patterns. It is less clear whether herbaceous perennials are roughly equivalent to one another, or whether seed-eating birds are roughly equivalent to seed-eating rodents because they consume the same resources. The exact meaning of "roughly equivalent" requires exploration within particular models. Indeed, the extent to which the assumption of equivalent species can be violated remains a thorny problem for theoretical and empirical research

Although models within the domain of this constitutive theory are often described as models of species diversity, they are more precisely models of species richness or species density (richness per unit area). Species richness is well defined. In contrast, there are many different definitions of diversity (Whittaker et al. 2001). All involve consideration of species richness, but also include the relative abundances or importances of the species. Some ecologists use the terms richness and diversity interchangeably. In almost all cases, discussions of "diversity gradients" are really discussions of "richness gradients." This is not a semantic argument, as gradients of different aspects of biodiversity (e.g., richness vs. evenness vs. diversity) can be quite different or even independent of each other (Stevens and Willig 2002, Chalcraft et al. 2009, Wilsey et al. 2005).

A THEORY OF ENVIRONMENTAL GRADIENTS OF SPECIES RICHNESS

Our theory rests on four propositions (Table 13.1), set within a conceptual framework (Figure 13.2). All models of gradients in species richness use the first two propositions, whereas only some include one or both of the last two. These propositions are not universal statements about
the world: we do not claim that all propositions hold under all circumstances. Rather, the propositions are statements about the structure of current models of ecological gradients.

Our four propositions are of different kinds. The first proposition is a definition, establishing the essence of an environmental gradient in abundance. The second proposition encompasses several mechanisms that can be derived from first principles, each of which assures that the environmental gradient in abundance is also a gradient in richness. The third proposition is a description of a common empirical pattern, or is a general statement about ecological variation. The fourth proposition comprises a heterogeneous mix of mechanisms that derive from other domains and theories that influence the nature of environmental gradients in species richness.

The scale of the data or the model, including aspects of grain and extent, determine the particular mechanisms in effect for each of the propositions. Propositions one and two (Table 13.1) are functions of the extent, the range of environmental conditions encompassed by the data or being described by a model. Proposition three is a function of grain, the sizes of sampled patches or local communities. These scales are always determined by the biology of the species under consideration. A failure to recognize the scales within which particular mechanisms operate has led to a misapplication of a much-cited model (Wright 1983) that has been used to explain global species richness gradients (see below). Recognizing such misapplication is an example of how the process of theory formalization (as illustrated throughout this book) can provide critical insights and guide future research.

Gradients

Our first proposition is that variation characterizes a limiting environmental factor *X*, which affects variation in the number of individuals that can persist in a sample location of a given size, thus creating an environmental gradient in abundance. The abundance gradient exists in space and time, although the environmental factor need not be auto-correlated spatially or temporally. Models typically consider only one or the other aspect.

This proposition is part of all models of gradients in species richness, but it is often implicit. The environmental factor could be concentrations of one or more resources, or some condition such as stress or disturbance. For the purposes of our presentation, the exact mechanism creating the link between number of individuals and the factor(s) *X*, does not matter

361

and will differ for each particular situation. Importantly, not all environmental variation creates variation in numbers of individuals, thereby constraining the domain to which our theory applies. In particular, the modifier "limiting" implies that the value of X at a site determines, at least at equilibrium, the number of individuals present, N(X). Most models assume that the system has approached some sort of long-term behavior, i.e., an asymptote or a dynamic equilibrium.

Careful consideration of this proposition helps to clarify limitations regarding the scope of particular models. If variation in X leads to variation in N, then we can write N(X) as a function predicting the equilibrial or asymptotic number of individuals. This requires that the model be general – it predicts the long-term number of individuals that can persist at a particular level of X, not the number or identities of species found at a particular instance of X. The individuals are therefore assumed to be identical in key ecological respects, such as physical traits like body size or demographic traits, insofar as these reflect resource use. This assumption holds only for those key characteristics. For example, models that invoke niche partitioning (Hutchinson 1959; Schoener 1974; Chesson and Huntly 1988; Leibold 1995; Rosenzweig 1995; Chase and Leibold 2003; Kelly and Bowler 2005; Chase Chapter 5) assume that species are equal only in their resource use.

While the mathematics in the current literature use the equality assumption, the models are universally interpreted as applying to cases where the species under consideration are roughly similar but not identical. Under this relaxed assumption, species must be sufficiently similar to one another that, to a rough approximation, individuals in one species require the same amount of resource, or respond in a similar fashion to some condition, as do those of any other species in the species pool. This could easily be generalized to allow for an equivalence among species (e.g., 1 individual of species A equals 1.7 individuals of species B). Thus, a model might describe gradients in the richness of granivorous rodents, but cannot be expected to describe gradients for all vertebrates or even all mammals. It might be reasonable to develop a model that describes a richness gradient for herbaceous C₃ plants with respect to some environmental characteristic. But since "all plants" includes organisms with profoundly different metabolic pathways that span 7 or 8 orders of magnitude in dry weight, a single model likely will not describe the gradient in richness of all plants.

Model construction is simple enough when the environmental factor is a single resource, but becomes more complicated if multiple resources govern variation in abundance. If the same resource is limiting for all species at each location or time, the minimum (limiting) resource dominates (Liebig's Law of the Minimum, Sprengel 1839; van der Ploeg et al. 1999). If two or more resources (say, water and light availability for plants) are both limiting (either simultaneously or each one limiting at different times or sites), the combination can be quantified by the vector \vec{X} . If the resources affect abundance additively, then the vector \vec{X} can be treated as a single resource, a linear combination of the multiple resources.

Thus, gradients can be grouped into two general classes. In the first class are systems constrained by a single factor X or a combination of factors \vec{X} , and $N(\vec{X})$ increases monotonically with the linear combination X_c . In the second class are systems in which one or more pairs of limiting factors are negatively correlated. The negative correlation could be intrinsic (e.g., as soil water content increases, oxygen levels in the soil must decrease), or could be extrinsic to the factors themselves.

Our description of the first class as following a single constraint needs further explanation because circumstances can be more complicated. Multiple resources can interact, so that more than one is limiting at a particular time or place (Gleeson and Tilman 1992). In principle, the only difference is that instead of the number of individuals $[N(\vec{X})]$ being a curve, it will be a surface with the number of dimensions equal to the number of factors that are limiting at some point. Consider a system in which two resources – each limiting at different concentrations – vary in a nonlinear fashion with respect to each other. Then it is possible for the system to switch between limitation by factor 1 and factor 2 more than once. Such is not a problem for the models, but in practice it may not be easy to determine what is limiting at each particular location. Indeed, most empirical studies do not address this concern. In practice, most studies examine only single factors (Scheiner and Willig 2005).

Finally, when gradients involve tradeoffs (proposition 4), it is generally useful to separately consider two different aspects of the environmental factors, one relating to resource attributes and the other to stress attributes. For example, consider communities of herbaceous plants arrayed from upland to wetland. Water may be a limiting resource at the higher elevations, but a stressor at lower elevations. It can be useful to model the responses of richness with the associated attributes as two separate factors because the biological responses to water as a resource and water as a stressor are different. Many models assume that $N(\vec{X})$ is a linear function, although the critical nature of this assumption has not been explored in a comprehensive fashion. When a single factor is limiting, the important assumption is that N(X) is monotonic. It is often possible to select a transformation (e.g., the log function) to linearize a monotonic pattern. However, multiple limiting factors that act singly or interactively create complications: if the factors are interactive, $N(\vec{X})$ may not be monotonic, and there may not be a transformation that will linearize the function.

One common instance of multiple factors is when richness is regulated by bottom-up and top-down interactions. If X is a resource (such as prey items) or an abiotic stressor, regulation is bottom-up; if X is predation, regulation is top-down. Because both top-down and bottom-up regulation occur in many systems, many models consider both kinds of factors in producing gradients of richness.

Individuals \propto Species

The proposition that the number of species increases with the number of individuals was developed by Fisher et al. (1943) and Preston (1962a). It has been termed the "More Individuals Hypothesis" (Srivastava and Lawton 1998), although it is not necessarily a hypothesis; under random placement it is a simple sampling relationship, but if extinction or speciation mechanisms are involved in creating the richness gradient, it is indeed a hypothesis. Three mechanisms can account for this pattern: random placement, local extinction, and speciation. Random placement and local extinction are modeled using similar mathematical constructions but are distinct in their biological causations. Moreover, they operate at somewhat different scales of time and space. Given an environmental gradient in the number of individuals, each of these mechanisms can lead to an environmental gradient in the number of species. Indeed, all models of species richness gradients invoke at least one of these mechanisms, at least implicitly. In many cases, the models focus on causes of gradients in the number of individuals and assume a mechanism whereby more individuals give rise to more species.

Random placement (also called passive sampling) refers to the movement of individuals among patches or communities. It creates a relationship between the number of individuals and the number of species if local species richness is determined by random sampling of individuals from a regional species pool (Coleman 1981; Coleman et al. 1982) or metacommunity (Hubbell 2001). Biologically, this occurs as individuals move independently but tend to concentrate in areas of greatest resource or least stress. The assumptions behind random placement models are thus identical to those leading to an "ideal free distribution," an idea that has played a critical role in behavioral ecology (Fretwell and Lucas 1970; Sih Chapter 4). As the number of individuals in a local area increases, the number of species should increase because the likelihood of including a rare species increases due to chance. Under this model the species identity of each individual is random, but the number of individuals in a local assemblage is not – it is given by the function N(X). That the relationship between local and regional richness is positive and monotonic does not depend on the abundance distribution locally or regionally, although those distributions determine the exact form of the relationship.

Local extinction is the mechanism invoked by Preston (1962a; 1962b) and highlighted most often as part of the MacArthur and Wilson (1967) theory of island biogeography (Sax and Gaines Chapter 10). This mechanism assumes that a local population will persist only above some minimum abundance. If an area holds more individuals, more populations can attain species-specific minimum viable sizes. Although details can differ about the exact form of the relationship between numbers of individuals and numbers of species, the core assumption is simply that the relationship is positive and monotonic.

Both random placement and local extinction share some similarities: both operate through a balance between the entry of individuals into a site and their departure – by movement under random placement, and by death in extinction. That entry and exit could occur within the lifetimes of individuals through movement, or it could occur across generations through colonization and extinction. For convenience we divide this continuum into an individual-level mechanism (random placement) and a population-level mechanism (local extinction). Mathematically, they can be treated as equivalent for an equilibrial theory, although models will differ in detail depending on the particular set of species under consideration. In general, random sampling operates at local to landscape scales over short time periods, whereas local extinction operates at landscape to regional scales over longer time periods, with the exact meaning of these differences determined by the species' biology.

This distinction between individual- and population-level processes can define the domain of a particular model. Previously, we asked whether seed-eating birds and seed-eating rodents are roughly equivalent. If the abundance of birds at sites is determined by the movement of individuals (e.g., Coleman et al. 1982), while the abundance of rodents is determined by

population growth and extinction, then application of existing models to, say, richness gradients of granivorous vertebrates would be misleading. Because the entry and departure processes are quite different for granivorous birds and rodents, involving different parameter values, one would need to model the richness of the two groups separately. If interactions between the two groups could be ignored, then the predicted richness of the combined taxa would simply be the sum of the two predicted richnesses. However, if granivorous birds and rodents interact, studying the richness of both would require modeling their interactions as well.

Speciation operates at scales of time and space that are much greater than that of random placement or local extinction. It assumes a positive relationship between the number of individuals and the net rate of speciation (i.e., speciation minus extinction; VanderMeulen et al. 2001). This mechanism most appropriately deals with species richness patterns at large spatial scales and may provide an explanation for the richness of the regional species pool.

All extant models of richness gradients make another important assumption: *they do not consider species interactions*. To see this, consider a gradient model derived from one of the versions of neutral theory (Hubbell 2001; Chave 2004; Etienne and Olff 2004; Volkov et al. 2005). A model using either random placement or local extinction can logically find the expected number of species at any location along the gradient, using only propositions 1 and 2 (Table 13.1). Now consider a gradient model that concerns niche partitioning, using character displacement, microhabitat variation, or temporal niches (e.g., Hutchinson 1959; Schoener 1974; Chesson and Huntly 1988; Leibold 1995; Rosenzweig 1995; Chase and Leibold 2003; Kelly and Bowler 2005). This model can also find the expected number of species, given one of the following assumptions:

- Local assemblages generally have reached persistent states. By persistent we include the textbook equilibria of Lotka-Volterra competition models, as well as the more complex kinds of persistence possible with multiple species and nonlinear interactions (Armstrong and McGehee 1980). This assumption implies that local population and community dynamics dominate, so logically this assumption might hold under a local extinction model, but not under a random placement model.
- The regional species pool contains only those species that can coexist with one another. This assumption could hold under random placement, which implies nothing

about the long-term persistence of competitors. It could also hold under the (rather unlikely) assumption that the species in the regional pool have all coevolved to coexist.

• Competitive coexistence of particular species in a given sample unit is not necessarily guaranteed, but sampling processes still guarantee that on average there will be *S*[*N*(*X*)] species at *X* resource level.

Having said this, we hasten to add that it is not logically necessary for community theory (Holt Chapter 7; Leibold Chapter 8; Pickett et al. Chapter 9) to be external to gradient theory – just that this is presently the case. It is certainly possible for local community dynamics to interact with the factors determining the existence of a richness gradient; addressing this possibility is an open theoretical question.

Mean ∝ *Variance*

The proposition that the mean and the variance of environmental characteristics are related positively is based, in part, on the recognition that most environmental factors are bounded by zero (i.e., have a theoretical minimum). Such a bound can lead to a positive relationship between the mean and variance, although such a relationship need not exist empirically. If the magnitude of an environmental factor is zero or close to zero, then perforce the variance initially will increase as the mean increases. A continued rise in the mean allows for the possible continued rise in the variance, unless an upper bound also exists (e. g., water saturation of soil). Thus, this proposition is limited to those environmental variables that have a lower but not an upper bound within the range of environmental conditions of the gradient. If the upper bounds on a limiting environmental factor also restrict the number of individuals, the theory as described in this chapter can be applied to that part of the environmental gradient where the variance does increase with the mean.

This mean-variance relationship is invoked in models that focus on patch dynamics (e.g., Abrams 1988). More specifically, species richness is measured in some area within which there are multiple patches. For some models, heterogeneity is generated by interactions among individuals (e.g., Tilman 1982; Huston 1994; Currie et al. 2004). Most commonly, the invoked mean-variance relationship is spatial (e.g., wet vs dry, good vs bad). A meta-analysis (Lundholm 2009) found that plant species richness or diversity frequently increases with spatial heterogeneity, but does not always do so. Some models invoke temporal heterogeneity,

considering specialization on different year-types as a mechanism that promotes coexistence of multiple species (Chesson and Huntly 1988; Rosenzweig 1995; Kelly and Bowler 2005). Regardless, if species specialize on combinations of environmental characteristics that occur in patches in which they can out-compete other species, then richness should increase as the number of patches (i.e., heterogeneity) in an area increases.

This proposition is explicitly scale-dependent (Lundholm 2009) as it deals with changes in variation within the grain of a particular model, the unit for which richness is measured. This grain is always dependent on the biology of the species under consideration. Thus proposition 2 also contains a hidden assumption that the species are equivalent in their use of space or time. At the lower end, the minimal grain size is that needed to hold one individual. At the upper end, the maximal grain size is such that all possible heterogeneity or habitat types are encompassed within a single grain.

The form of the relationship between mean patch characteristics and their variance is related to theories of species-area relationships (SARs; Figure 13.1). SARs are determined by a variety of factors: more individuals are contained in larger areas, and environmental heterogeneity increases with greater area. Clearly, models of SARs share many features with models of species richness gradients. Models of SARs are currently being developed and debated (e.g., Scheiner 2003; Tjørve 2003; Maddux 2004; Ostling et al. 2004; Adler et al. 2005; Fridley et al. 2006; Scheiner 2009; Chiarucci et al. submitted). Thus, we postpone any attempt to develop formal models of SARs specific to the context of species richness gradients until the more general forms of those models have been resolved more thoroughly.

Trade-offs and hump-shaped curves

Many models of environmental gradients in species richness posit that a trade-off leads to a hump-shaped pattern, with the maximum value of richness at some intermediate point along the axis of an environmental factor. The models differ with regard to the nature of the invoked trade-off. Nonetheless, they share the basic proposition that a change in the sign of the slope arises as a consequence of two mechanisms acting in concert but in an opposite fashion on each species. Commonly invoked trade-offs are competitive ability versus a variety of other abilities (e.g., stress tolerance, colonizing ability). The trade-offs that matter in a particular instance depend on

the species and type of environmental variation, including its scale (grain and extent) in time and space.

Trade-offs may occur at different levels in the biological hierarchy. For example, the trade-off may involve the characteristics of individuals, such as competitive ability versus stress tolerance (Grime 1973). In other cases, the trade-off may involve the characteristics of populations, such as the intensity of interspecific competition versus the intensity of predation (Oksanen et al. 1981). In yet other cases, the trade-off may involve characteristics of species, such as speciation rates versus extinction rates (VanderMeulen et al. 2001). Scheiner and Willig (2005, Table 1) listed 17 different models of species richness gradients. In the conceptual scheme presented here (Figure 13.2), we treat mechanisms that operate at the same level (i.e., individual, population or species) as mathematically equivalent. In doing so, we can unify some of those models, reducing the list of models from 17 to 9 (Table 13.2).

The maximum (or minimum) point in the curve describing an environmental gradient in species richness arises because of a change in the relative importance of factors that control the number of individuals. This trade-off can be conceptualized as environmental variation in each of two factors that are negatively correlated. Along one portion of the environmental axis, the first factor limits the number of individuals; at some point a second factor becomes limiting. This switch results in the number of individuals increasing along one portion of the environmental axis and decreasing along another. For many models, this shift in importance is controlled by inherent properties of species. For example, Tilman (1988) theorized that in terrestrial plant communities increasing nitrogen availability causes an increase in numbers of individuals, until plant density is great enough that light becomes limiting and numbers of individuals begins to decrease. Although trade-offs are invoked in models that produce a hump-shaped pattern, the mechanism can explain U-shaped patterns as well (Scheiner and Willig 2005).

In many models, the interacting mechanisms that determine the number of individuals are not stated explicitly. Similarly, the unique contributions of each mechanism to total abundance are rarely quantified with respect to variation in the environmental factors. As a result, the mechanistic trade-off is neither emphasized in conceptual models, nor detailed in quantitative models. The absence of mathematical or logical rigor enhances the likelihood that such concealment persists, diminishing an appreciation for the similarities of form that the details obscure. For example, various models posit trade-offs between competition for different resources (e.g., Tilman 1982; 1988; Huston 1994) or competition vs resistance to predation/herbivory (Leibold 1996; 1999). Such models all have a similar mathematical form, but this similarity is not apparent until they are placed within a single framework.

Arguments for and against particular models often boil down to a personal preference for one trade-off versus another. We take a more catholic position by not advocating any one in particular. Rather, we embrace all of them as theoretical possibilities, although it remains to be seen if some trade-offs are more common than others. Perhaps most critically, the posited mechanisms often are not mutually exclusive. Trade-offs may simultaneously exist between competition for two different resources and herbivory, for example. As with multiple environmental factors, it may be possible to model such multiple trade-offs as an additive pair of trade-offs. Otherwise, more complex models will be needed.

Although a specific trade-off may exist for a particular set of species, we should not expect the same trade-off to be ubiquitous for all species in a guild, trophic level, or community, thus limiting the scope of any particular model. It is possible that more closely related species will share a trade-off, whereas more distantly related taxa will have different constraints, but this should not be assumed (Losos 2008). Thus, the type and form of trade-offs sets another boundary on the conditions under which individuals of different species must be roughly equivalent. It is not known how rough this equivalence can be and still be consistent with the underlying models.

Relationship to the theory of ecology

The four propositions of the theory of environmental gradients of species richness (Table 13.1) derive from the fundamental principles of the theory of ecology (Table 1.3). Proposition 1 is a consequence of principles 4 or 5, depending on the nature of the environmental factor(s). The finite nature of resources (principle 5) creates the constraint that allows one or more resources to be limiting. The heterogeneity of environmental characteristics in space or time (principle 4) creates the potential for variation in resources or stressors. Environmental heterogeneity in time leads to the potential for variation among patches in the rate of disturbance. Proposition 2 is a consequence of principles 1, 2 or 7. The process of random sampling is one mechanism that creates the heterogeneous distribution of organisms (principle 1). Immigration-extinction balance comes about through the combination of processes that lead to heterogeneous distributions or organisms and species' interactions (principle 2). Speciation is a suite of processes that derive

from principle 7. Proposition 3 is a direct manifestation of environmental heterogeneity (principle 4). Finally, the trade-offs embodied in proposition 4 derive from principles 6 and 7.

APPLYING THE THEORY: THE ENERGY MODEL

To see how the constitutive theory relates to current models of productivity-diversity relationships, consider the energy model (Connell and Orias 1964; Wright 1983). We focus on this model for two reasons: (1) it has been very influential [we found 343 citations of Wright (1983) in the Web of Science database on February 16, 2009], and (2) it is one of the few that is written in explicit mathematical form. We follow the formal presentation of Wright (1983), which is couched in terms of the relationship of species richness and area. The model predicts the number of species in a sampling unit (Wright thought of these as islands) as a function of local energy availability. Wright considered energy input per unit area to be fixed so that his model predicts the consequences of variation in area on species richness. By contrast, the models considered in this chapter examine the consequences of variation in environmental resources or stressors among different locations while holding area constant. Thus, our explication of this model does not include terms for area as in Wright (1983).

Wright's model is $S = a (E \rho/m)^{z}$ where *E* is the amount of energy locally available for biosynthesis, ρ is an empirical constant for a given set of species describing the number of individuals supported per unit of available energy, and *m* is the population size of the smallest extant population. The terms *a* and *z* are empirical constants estimated from the data, although as we shall see, *a*, *z*, and *m* appear in this model because of some strong assumptions. We now examine how this model relates to our propositions, and consider some consequences of its assumptions.

Existence of a gradient and its consequences

Rewriting Wright's model in our more general terms, we begin with $N(X) = X\rho$. The use of the common term ρ means that the model describes richness when derived from a set of roughly similar species, as our explication of the general theory suggests it must. Wright further posited that species richness increases with decreasing latitude because available energy increases, a contention still advanced by many (e.g., Mittelbach et al. 2001; Hawkins et al. 2003). This is a

sensible model only if individuals of all species along the latitudinal gradient require about the same level of resources, which is certainly not true.

From individuals to species

Proposition 2 posits that more individuals lead to more species, and in Wright's model, most of the action is in proposition 2. To model species richness as a function of *X*, we need to model S = f[N(X)], where *S* is the number of species present and *f* is some function. Wright's choice of *f* is $f[N] = a(N/m)^z$, which comes from Preston (1962a). This equation, with $N = X\rho$ as above, produces a positive monotonic relationship between the resource *X* and species richness *S*. The exact shape of the relationship depends on *a*, *z*, and *m*. The model requires the first two propositions (Table 13.1), and nothing more. The energy model does not attempt to explain the source of available energy or its relationship to climate, which is the domain of other theories (e.g., O'Brien et al. 2000).

The Wright and Preston models rely on the local extinction mechanism, delimiting the temporal and spatial scales for which the model makes predictions. In particular, this model makes predictions about the equilibrial number of species at a location with resource concentration *X*, when individuals are drawn from a fixed regional pool of species. The model should not be interpreted as making predictions about variation in species richness over large spatial extents (e.g., across continents) because such variation cannot result from local extinctions from a single species pool. Such large-scale gradients must involve (at least) several species pools, and likely involve speciation processes as well. Thus, by its implementation of both Propositions 1 and 2, Wright's model has a far more limited interpretation than stated by Wright or many subsequent authors (e.g., Currie 1991; Monkkonen and Viro 1997; O'Brien 1998; Gaston 2000; Allen et al. 2002; Currie et al. 2004). Although some found apparently good fits of the model for continental-to-global scale data, because those data represent an inappropriate spatial domain, it is illogical to assign meaning to estimated parameters in terms of the Wright model.

Although the power-law function used by Preston and Wright is simple and familiar to several generations of ecologists, its derivation in this case rests on a complex and rather narrow argument concerning the distribution of species abundances and how population size relates to extinction probability. In particular, the power-law function depends on Preston's assumption

that species-abundance curves are described by a form of the lognormal distribution that he termed canonical.

The division by m – not explained by Wright other than his citation of Preston (1962a) – seems odd. It is natural, albeit wrong, to assume that this division (*N/m*) is aimed at calculating the maximum number of species. Under Preston's (1962a) canonical lognormal distribution, one specifies the shape of the species-abundance distribution with any two of three quantities: the total number of species, the standard deviation of the lognormal distribution, and the number of species in the modal octave. The quantity m is required to specify the position of this distribution along the horizontal axis (the log₂ of abundance). Preston (1962a) calls m the size of the smallest population, but he also calls this a "tentative" definition (Preston 1962a, p. 190), and notes "in practice that m is less, even appreciably less, than unity, and the temporary interpretation we have given [as the size of the smallest population] then has no meaning." In other words, m is just a parameter that defines the location of the species-abundance distribution, in the same sense that statisticians speak of the mean as characterizing the location of the normal distribution.

Regardless of whether *m* is the size of the smallest population or an empirically estimated parameter, Wright's use of *m*, *a*, and *z* to define his model links it intimately to the somewhat arbitrary assumptions of Preston's canonical lognormal distribution. Despite numerous criticisms of aspects of Preston's work (e.g., Pielou 1969; May 1975; Williamson and Gaston 2005), it has had remarkable staying power in the ecological literature. This is remarkable because neither Preston nor subsequent researchers have linked the canonical lognormal to any underlying mechanisms. Preston himself (1962a) made it clear that he had none in mind. Unfortunately, ecologists are sometimes satisfied with curve-fitting exercises without concern with the underlying mechanisms. Such exercises teach us nothing beyond the narrow lesson that the particular data set is well-described by a particular curve, providing only a phenomenological description.

Most ecologists, trying to justify the use of the lognormal for species abundance distributions, do so with a vague and incorrect reference to the central limit theorem (Williamson and Gaston 2005). The central limit theorem predicts that each species' abundance will be lognormally distributed over time; unless the abundances are independently and identically distributed among species (i.e., the species are equivalent), this does not lead to a jointly lognormal distribution of abundances at a given time. If the species are different from one another (i.e., they have a different means and variances of abundance), the joint distribution of abundances at a given time will not be lognormal. Šizling et al. (2009) proposed a more satisfying (and rigorous) explanation as to why species abundance distributions are often similar to the lognormal. Their derivation requires only that the abundance distribution be based on the combination of abundances in many non-overlapping subplots.

None of this implies that the Wright energy model is wrong in some sense; rather, its basis is weaker than one might hope (given its influence), as it depends on the phenomenological assumption that species abundances are given by Preston's canonical lognormal. Other mechanisms could be invoked that yield the same qualitative relationship while differing in details (e.g., Hubbell 2001). Pueyo et al. (2007) showed that an infinite number of models varying between strict neutrality (all species identical) and strict idiosyncrasy (all species unique) can generate identical abundance patterns. Also, while Wright's model has been interpreted as a predictor of continental to global patterns, it cannot logically do so as it is restricted to a set of roughly equivalent species (implementation of proposition 1) in a single regional pool (implementation of proposition 2).

The converse is also true. Many studies have shown a positive relationship between energy and species richness on a continental to global scale (e.g., Field et al. 2005; Rodriguez et al. 2005; Buckley and Jetz 2007; Davies et al. 2007; Kalmar and Currie 2007; Kreft and Jetz 2007; Woodward and Kelly 2008) and this has often been taken as support for Wright's model. However, those studies do not attempt to directly parameterize Wright's model and test whether the model is accurately predicting those relationships. Instead, we merely have a qualitative agreement between various empirical relationships and one particular model. Our dissection of that model suggests that it cannot be used as an explanation for those relationships because the mechanisms underlying that model operate at different scales. Given the generality of the observed relationships, further work is necessary to connect the mechanisms operating at that scale.

Our explication of Wright's model suggests that it must be interpreted on a regional spatial scale with species that are roughly equivalent. That does not preclude the possibility that one might find that the model provides a good fit to data from much larger spatial or taxonomic scales. Indeed, if a model like Wright's provides a good prediction of the number of species, given a level of resources, it may be useful to managers even if its assumptions are violated severely. The only problem here is with interpretation: a good fit of a model to data that violate its assumptions cannot be interpreted as support for the concepts embodied in the model, but only as a useful description of data. Prediction and understanding are not always on the same footing.

PROSPECTS

Model development

Theory unification is an iterative process that includes recognition of similarities among ostensibly competing models, development of a common framework, and construction of new overarching models within that framework. Additional effort is needed in domains, such as the one we consider, in which many of the models are verbal and even the analytic models have not been examined deeply. We are encouraged that our refinement of the conceptual framework (Figure 13.2) has led to further model unification (i.e., reducing the number of models from 17 to 9).

This is a step forward in model unification not simply because it reduces the number of models, but because it reveals their common bases, and because it makes it points to some additional models that have not yet been studied (Figure 13.2). The reduction in the number of models is a consequence of recognizing that the 14 different forms for proposition 4 listed in Scheiner and Willig (2005, Table 1) can be usefully placed in three categories: tradeoffs operating at the levels of individuals, populations, and species. For example, using the model numbers from Scheiner and Willig (2005, Table 1), we now treat models 3, 7, and 10 – all of which assume that the gradient (proposition 1) is productivity or stress, the number of species (proposition 2) is generated by local extinction processes, heterogeneity (proposition 3) occurs over space, and tradeoffs (proposition 4) occur at the levels of individual characteristics. Similar reasoning leads to the combining of other models.

Our approach has been to start with the simplest formal model, the Wright energy model, and carefully examine its assumptions and limitations. Building a new, general and useful model that avoids the limitations described above is a challenge. The first limitation – restriction to a set of species with roughly equivalent requirements – is a hurdle only if one hopes to develop a model that explains richness in general. To the extent that progress can be made studying richness gradients of given taxa or guilds, there is no limitation. If interest lies in explaining

more general gradients, however, it is not logically possible to follow the approach of first calculating the number of individuals (proposition 1) and then using a sampling argument (proposition 2) – either random placement or local extinction – to predict the number of species.

A more general model must incorporate the rules by which metacommunities are formed (Leibold Chapter 8). In other words, such a model would need our four propositions as well as propositions involving the way in which interactions among species determine and the numbers of individuals and species. One might argue that this is precisely what Preston (1948; 1962a; 1962b) attempted, but this is not the case. Preston's argument was couched entirely in terms of single species. We know of no persuasive models that jointly predict the numbers of species and the population sizes of multiple species.

How might we avoid the second limitation, being wed to a set of arbitrary assumptions necessary to go from N(X), the number of individuals, to S = f[N(X)], the number of species? Numerous models of species abundance distributions arise from quite different assumptions (Fisher et al. 1943; Preston 1948; Zipf 1965; Kempton and Taylor 1974; Pielou 1975; Mandelbrot 1977; Engen and Lande 1996; Engen 2001; Hubbell 2001; Dewdney 2003; Lande et al. 2003; Williamson and Gaston 2005). At this point there is no basis for concluding that any particular model is either logically best or empirically most supported by available data. In the absence of such a model, assumptions about the form of S = f[N(X)] are arbitrary. This does not necessarily mean that more progress in gradient theory must await developments in the theory of species abundance distributions. It is possible to make progress by using a number of different species abundance distributions and asking how the choice of distribution affects the model predictions about richness gradients. Many gradient models may be robust to such choices.

Further work is needed to relax the assumption of species equivalence. For example, for the random placement mechanism, one could substitute a distribution of body mass frequencies for the constant ρ . Such a model would still assume that the shape of the distribution is the same for all sites, but that is a much weaker assumption.

Linking models to data

More challenging than model development is linking models to data. Even for a model as simple as Wright's energy model, which does not invoke trade-offs or spatial structure, the information necessary to estimate all of the parameters does not exist, as far as we are aware. When confronted by such challenges, ecologists often respond by questioning the utility of the model. Our reply is two-fold. First, formalizing models makes data requirements clearer. Although many data have been gathered in the context of studying richness gradients (Mittelbach et al. 2001, Gillman and Wright 2006), those studies have not been guided by theory, thus the disconnect between the data and the models. For example, few studies collect data on richness, abundance, and the environmental variables thought to determine richness and abundance. It may be that sufficient data exist for some systems and the challenge is to discover and assemble those data.

Second, only models can provide quantitative predictions. Enough may be known about processes such as herbivory or competition to permit a sufficiently constrained state-space within which a model can be explored. Given the growing urgency of understanding global change, these models, with their general parameters, may have to do while we work to collect more data. For example, our demonstration that as a mechanistic model, the Wright energy model should be restricted to local or regional gradients and limited sets of taxa or guilds suggests that it should not be combined with global change models to predict changes in global species richness, or that any such model should be sharply delimited in its taxonomic or ecological scope (e.g., Field et al. 2005). Obviously, Wright's model can still be used on these scales as a phenomenological model, so long as interpretation of the fit and parameter estimates is appropriately restricted. Similar hidden limitations may be discovered as we explore the details of other models.

Linkages to other constitutive theories

The theory of gradients of species richness has direct linkages to many of the other constitutive theories presented in this book. Geographic gradients (Colwell Chapter 14) concern spatial gradients only; the models considered here may have a spatial component, but typically do not. Not surprisingly, the theories shares points of contact concerning the multiplicity of causes that determine gradients and how variation in species ranges along a gradient determine the form of the species richness relationship. Island biogeography theory (Sax and Gaines Chapter 10) is another one with shared mechanisms concerning immigration and extinction (Table 10.1, propositions 1, 4 and 6). The Wright energy model was first developed within the context of island biogeography theory as a way of explaining the relationship between area and species richness. As we have discussed, metacommunity theory (Leibold Chapter 8) may provide

important tools for linking species abundance and species richness. Similarly, in order to formalize models that invoke trade-offs in competition or predation/herbivory will require examining niche theory (Chase Chapter 5) and enemy-victim theory (Holt Chapter 7). Thus, the entire processes of theory formalization represented by this book will be an important guide and useful tool for further model development.

Acknowledgements

The evolution of ideas presented in the chapter arose in part as a consequence of a workshop held at the University of Connecticut and supported by the Center for Environmental Science and Engineering. We thank Nancy Huntly and Dov Sax for their comments and suggestions. Support to MRW was provided by the Center for Environmental Sciences and Engineering at the University of Connecticut, as well as by National Science Foundation grant DEB-0218039. GAF was partially supported by National Science Foundation grant DEB-0614468 and by a sabbatical from the University of South Florida. This manuscript is based on work done by SMS while serving at the U.S. National Science Foundation and on sabbatical at both the Center for Environmental Sciences and Engineering at the University of South Florida. The views expressed in this paper do not necessarily reflect those of the National Science Foundation or the United States Government. Literature Cited

- Abrams, P. A. 1988. Resource productivity-consumer species diversity: simple models of competition in spatially heterogeneous environments. Ecology 69:1418-1433.
- Adler, P. B., E. P. White, W. K. Lauenroth, D. M. Kaufman, A. Rassweiler, and J. A. Rusak. 2005. Evidence for a general species-time-area relationship. Ecology 86:2032-2039.
- Allen, A. P., J. H. Brown, and J. F. Gillooly. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545-1548.
- Armstrong, R. A., and R. McGehee. 1980. Competitive exclusion. American Naturalist 115:151-170.
- Buckley, L. B., and W. Jetz. 2007. Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B-Biological Sciences 274:1167-1173.
- Chase, J. M., and M. A. Leibold. 2003, Ecological Niches: Linking Classical and Contemporary Approaches. Chicago, University of Chicago Press.
- Chalcraft, D. R., B. J. Wilsey, C. Bowles and M. R. Willig. 2009. The relationship between productivity and multiple aspects of biodiversity in grassland communities. *Biodiversity and Conservation* 18:91-104.
- Chave, J. 2004. Neutral theory and community ecology. Ecology Letters 7:241-253.
- Chesson, P. L., and N. Huntly. 1988. Community consequences of life history traits in a variable environment. Annales Zoologici Fennici 25:5-16.
- Chiarucci, A., G. Bacaro, D. Rocchini, C. Ricotta, M. W. Palmer, and S. M. Scheiner. (submitted). Spatially Constrained Rarefaction: incorporating the autocorrelated structure of biological communities in sample-based rarefaction. Community Ecology.
- Coleman, B. D. 1981. On random placement and species-area relations. Mathematical Biosciences 54:191-215.
- Coleman, B. D., M. A. Mares, M. R. Willig, and Y.-H. Hsieh. 1982. Randomness, area, and species richness. Ecology 63:1121-1133.
- Connell, J. H., and E. Orias. 1964. The ecological regulation of species diversity. American Naturalist 98:399-414.
- Currie, D. J. 1991. Energy and large-scale patterns of animal-and plant-species richness. American Naturalist 137:27-49.

- Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J.-F. Guégan, B. A. Hawkins, D. M. Kaufman et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters 7:1121-1134.
- Davies, R. G., C. D. L. Orme, D. Storch, V. A. Olson, G. H. Thomas, S. G. Ross, T. S. Ding et al. 2007. Topography, energy and the global distribution of bird species richness.
 Proceedings of the Royal Society B-Biological Sciences 274:1189-1197.
- Denslow, J. S. 1980. Gap partitioning among tropical rainforest trees. Biotropica (Supplement) 12:47-55.
- Dewdney, A. K. 2003. The stochastic community and the logistic-J distribution. Acta Oecologia 24:221-229.
- Engen, S. 2001. A dynamic and spatial model with migration generating the log-Gaussian field of population densities. Mathematical Biosciences 173.
- Engen, S., and R. Lande. 1996. Population dynamic models generating species abundance distributions of the gamma type. Journal of Theoretical Biology 178:325-331.
- Etienne, R. S., and H. Olff. 2004. A novel genealogical approach to neutral biodiversity theory. Ecology Letters 7:170-175.
- Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: development and evaluation. Ecology 86:2263-2277.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. . Journal of Animal Ecology 12:42-58.
- Fridley, J. D., R. K. Peet, E. van der Maarel, and J. H. Willems. 2006. Integration of local and regional species-area relationships from space-time species accumulation. American Naturalist 168:133-143.
- Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405:220-227.
- Gillman, L. N., and S. D. Wright. 2006. The influence of productivity on the species richness of plants: a critical assessment. Ecology 87:1234-1243.
- Gleeson, S. K., and D. Tilman. 1992. Plant allocation and the multiple limitation hypothesis. The American Naturalist 139:1322-1343.
- Grace, J. B. 1999. The factors controlling species density in herbaceous plant communities: an assessment. Perspectives in Plant Ecology, Evolution and Systematics 2:1-28.

Grime, J. P. 1973. Competitive exclusion in herbaceous vegetation. Nature 242:344-347.

- —. 1979, Plant Strategies and Vegetation Processes. New York, John Wiley and Sons.
- Hawkins, B. A. 2001. Ecology's oldest pattern? Trends in Ecology and Evolution 16:470.
- Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J.-F. Guégan, D. M. Kaufman, J. T. Kerr et al. 2003. Energy, water and broad-scale geographic patterns of species richness. Ecology 84:3105-3117.
- Hubbell, S. P. 2001, The Unified Neutral Theory of Biodiversity and Biogeography: Princeton Monographs in Population Biology. Princeton, NJ, Princeton University Press.
- Huston, M., and T. Smith. 1987. Plant succession: life history and competition. American Naturalist 130:168-198.
- Huston, M. A. 1979. A general hypothesis of species diversity. American Naturalist 113:81-101.
- —. 1994, Biological Diversity: the coexistence of species in changing landscapes. Cambridge, UK, Cambridge University Press.
- Huston, M. A., and D. L. DeAngelis. 1994. Competition and coexistence: the effects of resource transport and supply rates. American Naturalist 144:954-977.
- Hutchinson, G. E. 1959. Homage to Sanda Rosalia or Why are there so many kinds of animals? American Naturalist 93:145-159.
- Kalmar, A., and D. J. Currie. 2007. A unified model of avian species richness on islands and continents. Ecology 88:1309-1321.
- Kelly, C. K., and M. G. Bowler. 2005. A new application of storage dynamics: Differential sensitivity, diffuse competition, and temporal niches. Ecology 86:1012-1022.
- Kempton, R. A., and L. R. Taylor. 1974. Log-series and lognormal parameters as diversity discriminates for the Lepidoptera. Journal of Animal Ecology 43:381-399.
- Kreft, H., and W. Jetz. 2007. Global patterns and determinants of vascular plant diversity. Proceedings of the National Academy of Sciences of the United States of America 104:5925-5930.
- Lande, R., S. Engen, and B.-E. Saether. 2003, Stochastic Population Dynamics in Ecology and Conservation. Oxford, Oxford University Press.
- Leibold, M. A. 1995. The niche concept revisited: mechanistic models and community context. Ecology 76:1371-1382.

- —. 1996. A graphical model of keystone predators in food webs: trophic regulation of abundance, incidence, and diversity patterns in communities. American Naturalist 147:784-812.
- . 1999. Biodiversity and nutrient enrichment in pond plankton communities. Evolutionary Ecology Research 1:73-95.
- Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity between species. Ecology Letters 11: 995-1007.
- Lundholm, J. T. 2009. Plant species diversity and environmental heterogeneity: spatial scale and competing hypotheses. Journal of Vegetation Science 20: 377-391.
- MacArthur, R. H., and E. O. Wilson. 1967, The Theory of Island Biogeography. Princeton, NJ, Princeton University Press.
- Maddux, R. D. 2004. Self-similarity and the species-area relationship. American Naturalist 163:616-626.
- Mandelbrot, B. B. 1977, Fractals, Fun, Chance and Dimension. New York, Freeman.
- May, R. M. 1975. Patterns of species abundance and diversity, Pages 81-120 *in* M. L. Cody, andJ. M. Diamond, eds. Ecology and evolution of communities. Cambridge, MA, BelknapPress of Harvard University Press.
- Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig et al. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381-2396.
- Monkkonen, M., and P. Viro. 1997. Taxonomic diversity of the terrestrial bird and mammal fauna in temperate and boreal biomes of the northern hemisphere. Journal of Biogeography 24:603-612.
- O'Brien, E., R. Field, and R. J. Whittaker. 2000. Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. Oikos 89:588-600.
- O'Brien, E. M. 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. Journal of Biogeography 25:379-398.
- Oksanen, J. 1996. Is the humped relationship between species richness and biomass an artefact due to plot size? Journal of Ecology 84:293-295.

- Oksanen, L., S. D. Fretwell, J. Arrunda, and P. Niemela. 1981. Exploitation ecosystems in gradients of primary productivity. American Naturalist 118:240-261.
- Ostling, A., J. Harte, J. L. Green, and A. P. Kinzig. 2004. Self-similarity, the power law form of the species-area relationship, and a probability rule: a reply to Maddux. American Naturalist 163:627-633.
- Pielou, E. C. 1969, Introduction to Mathematical Ecology. New York, Wiley.
- —. 1975, Ecological Diversity. New York, Wiley.
- Preston, F. W. 1948. The commonness and rarity of species. Ecology 29:254-283.
- —. 1962a. The canonical distribution of commonness and rarity: Part I. Ecology 43:182-215.
- —. 1962b. The canonical distribution of commonness and rarity: Part II. Ecology 43:410-432.
- Pueyo, S., F. He, and T. Zillio. 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecology Letters 10:1017-1028.
- Rodriguez, M. A., J. A. Belmontes, and B. A. Hawkins. 2005. Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. Acta Oecologica-International Journal of Ecology 28:65-70.
- Rosenzweig, M. L. 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. Science 171:385-387.
- —. 1995, Species Diversity in Space and Time. Cambridge, UK, Cambridge University Press.
- Rosenzweig, M. L., and Z. Abramsky. 1993. How are diversity and productivity related?, Pages 52-65 *in* R. E. Ricklefs, and D. Schluter, eds. Species Diversity in Ecological Communities. Chicago, University of Chicago Press.
- Schaffer, W. M. 1983. On the application of optimal control theory to the general life history problem. American Naturalist 121:418-431.
- Scheiner, S. M. 2003. Six types of species-area curves. Global Ecology and Biogeography 12:441-447.
- Scheiner, S. M. 2009. The terminology and use of species-area relationships: reply to Dengler. Journal of Biogeography (in press).
- Scheiner, S. M., and M. R. Willig. 2005. Developing unified theories in ecology as exemplified with diversity gradients. American Naturalist 166:458-469.
- Schoener, T. W. 1974. Resource partitioning in ecological communities. Science 185:27-39.

- Šizling, A. L., D. Storch, E. Šizlingová, J. Reif, and K. Gaston. 2009. Species abundance distribution results from a spatial analogy of central limit theorem. PNAS 106: 6691-6695.
- Sprengel, C. 1839, Die Lehre vom Dünger oder Beschreibung aller bei der Landwirthschaft Gebräuchlicher Vegetablilischer, Animalischer und Mineralischer Düngermaterialien, nebst Erklärung ihrer Wirkungsart. Leipzig.
- Srivastava, D. S., and J. H. Lawton. 1998. Why more productive sites have more species: an experimental test of theory using tree-hole communities. American Naturalist 152:510-529.
- Stevens, R. D., and M. R. Willig. 2002. Geographical ecology at the community level: perspectives on the diversity of New World bats. Ecology 83: 545-560.
- Stevens, M. H. H., and W. P. Carson. 1999. The significance of assemblage level thinning for species richness. Journal of Ecology 87:490-502.
- Tilman, D. 1982, Resource Competition and Community Structure. Princeton, NJ, Princeton University Press.
- Tilman, D., and S. Pacala. 1993. The maintenance of species richness in plant communities, Pages 13-25 in R. E. Ricklefs, and D. Schluter, eds. Species Diversity in Ecological Communities. Chicago, University of Chicago Press.
- Tjørve, E. 2003. Shapes and functions of species-area curves: a review of possible models. Journal of Biogeography 30:827-835.
- van der Ploeg, R. R., W. Böhm, and M. B. Kirkham. 1999. On the origin of the theory of mineral nutrition of plants and the law of the minimum. Soil Science Society of America Journal 63:1055-1062.
- VanderMeulen, M. A., A. J. Hudson, and S. M. Scheiner. 2001. Three evolutionary hypotheses for the hump-shaped productivity-diversity curve. Evolutionary Ecology Research 3:379-392.
- Volkov, I., J. R. Banavar, F. He, S. P. Hubbell, and A. Maritan. 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438:658-661.
- von Humboldt, A. 1808, Ansichten der Natur mit wissenschaftlichen Erlauterungen. Tübingen, Germany, J. G. Cotta.

- Waide, R. B., M. R. Willig, C. F. Steiner, G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday, and R. Parmenter. 1999. The relationship between productivity and species richness. Annual Review of Ecology and Systematics 30: 257-300.
- Whittaker, R. J., K. J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. Journal of Biogeography 28:453-470.
- Williamson, M., and K. J. Gaston. 2005. The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. Journal of Animal Ecology 74:409-422.
- Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics 34:273-309.
- Wilsey, B. J., D. R. Chalcraft, C. M. Bowles, and M. R. Willig. 2005. Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. Ecology 86: 1178-1184.
- Wollkind, D. J. 1976. Exploitation in three trophic levels: an extension allowing intraspecies carnivore interaction. American Naturalist 110:431-447.
- Woodward, F. I., and C. K. Kelly. 2008. Responses of global plant diversity capacity to changes in carbon dioxide concentration and climate. Ecology Letters 11:1229-1237
- Wright, D. H. 1983. Species-energy theory: an extension of species-area theory. Oikos 41:496-506.
- Zipf, G. K. 1965, Human Behavior and the Principle of Least Effort. New York, Hafner.

Table 13.1. The domain, background assumptions, and propositions that constitute the theory of species richness gradients. Propositions 1 and 2 are used by all models, whereas propositions 3 and 4 are only used by some.

Domain	Environmental gradients in species richness. The gradient can extend over very short spatial distances or be global, or it can extend over short or very long periods of time.
Assumptions	Systems are at equilibrium at some spatial or temporal scale. [most models] The species under consideration are roughly equivalent in their resource requirements, dispersal abilities, and extinction probabilities.
	Each species restricts itself more than it restricts other species. Local assemblages tend to be in persistent states. [local extinction models] The regional species pool contains only species that can coexist with one another [random placement models]

Propositions

- A gradient implies one or more limiting resources or conditions that differ in space or time.
- 8. In a uniform environment of fixed area, more individuals lead to more species.
- 9. Within an area of fixed size or a unit of time of fixed duration, the variance of an environmental factor increases with its mean.
- 10. All non-monotonic relationships require a trade-off in organismal, population, or species characteristics with respect to the environmental gradient.

Table 13.2. Models of diversity gradients and their components and mechanisms. Proposition 1: Type of gradient; Proposition 2: Mechanism linking the number of individuals and number of species; Proposition 3: Environmental heterogeneity; Proposition 4: Type of trade-off. Previous number(s) refers to models listed in Table 1 of Scheiner and Willig (2005).

NT					Previous	0
No.	Proposition I	Proposition 2	Proposition 3	Proposition 4	number(s)	Sources
	Productivity	Random		Individual		
1	or stress	placement	N/A	characteristics	1	Oksanen (1996), Stevens (1999)
	Productivity	Local				
2	or stress	extinction	N/A	N/A	5	Connell (1964), Wright (1983)
	Productivity	Local		Individual		
3	or stress	extinction	N/A	characteristics	11	Rosenzweig (1993), Tilman (1993)
	Productivity	Local		Individual		Tilman (, 1982 #60; , 1988 #61}, Abrams
4	or stress	extinction	Spatial	characteristics	3, 7, 10	(1988), Huston (1994), Leibold (1996; 1999)
	Productivity	Local		Individual		
5	or stress	extinction	Temporal	characteristics	4	Rosenzweig (1995)
	Productivity					
	or stress and	Local		Individual		Grime (1973; 1979), Huston (1979), Huston and
6	disturbance	extinction	Temporal	characteristics	6, 8	Smith (1987)

	Productivity	Local		Population		Rosenzweig (1971; 1995), Wollkind (1976),
7	or stress	extinction	N/A	characteristics	2,9	Oksanen et al. (1981)
	Productivity			Individual		Denslow (1980), Rosenzweig and Abramsky
8	or stress	Speciation	N/A	characteristics	12, 15	(1993), VanderMeulen et al. (2001)
	Productivity			Species		
9	or stress	Speciation	N/A	characteristics	13, 14	VanderMeulen et al. (2001)

Figure 13.1. Gradient theory and species-area theory both predict species richness (S) by fixing the value of the other axis (area and resource/stressor, respectively) and making landscape-level assumptions. (A) For gradient theory the relationship can take a variety of forms; we depict three here. For species-area theory the relationship is generally assumed to be monotonically increasing or asymptotic. (B) We do not yet know how to draw the surface connecting these graphs in general, because we do not know whether area can interact with the resource or stressor axis, and in most cases would need additional axes for landscape-level variables. The surface on the right shows what the resulting model would look like for a unimodal resource gradient in a uniform landscape without interactions between the axes.



Figure 13.2. A diagram indicating how the four propositions (Table 13.1) can be assembled into different models. The vertical dotted line separates the two propositions (1 and 2) that must be included in any model from the two (3 and 4) that are optional. Solid arrows indicate propositions that have been linked in at least one model. Not all possible combinations of linkages appear in current models. Of the 43 possible models based on unique combinations of linkages, only 9 have been developed to date. Dashed arrows indicate linkages that have not been made. Additional linkages could be developed between propositions 2 and 3, but are not included in the diagram for clarity. The absent linkages between propositions 2 and 4 may not exist because of incommensurate time scales; however, we do not preclude the development of such linkages.



Chapter 14: Biogeographical Gradient Theory Robert K. Colwell

THE EMERGENCE OF BIOGEOGRAPHICAL GRADIENT THEORY

The history and contemporary distribution of life on Earth on broad spatial scales has traditionally circumscribed the realm of biogeography. Early observers, beginning in Classical times and continuing through the 18th and 19th Centuries, began by describing where the most conspicuous species were found, accumulating knowledge of vertebrates, plants, and the larger and more showier arthropods on land, and of the more conspicuous, accessible and useful macro-organisms of the seas. Meanwhile, the descriptive geography of fossils launched the twin disciplines biostratigraphy and paleogeography. The same explorations that provided the foundation for descriptive biogeography, of both extinct and living species, provided the material for taxonomists to describe species and to classify them in increasingly "natural" taxa. Knowledge of the taxonomy and geographical distribution of conspicuous species reached a critical stage in the mid-19th Century, making possible the independent discovery by Wallace and Darwin of the theories of natural selection and speciation by isolation, which transformed all of biology.

In spite of insightful speculation by 19th Century observers (e.g. von Humboldt 1807, Wallace 1878) and the emergence early in the 20th Century of the landscape-scale work of early plant ecologists that would eventually lay the foundations of community ecology, there was no formal theory of the geography of species and biotas on broad spatial scales. It took nearly another century of accumulated biogeographical knowledge to set the stage for the emergence of biogeographical theory, which many would trace to two pinnacles of the brief career of R. H. MacArthur: his collaboration with E. O. Wilson in *The Theory Of Island Biogeography* (MacArthur and Wilson 1967), and his extension of ecological theory to broader spatial scales in the publication of *Geographical Ecology* (MacArthur 1972).

Biogeographical gradients—spatial patterns in the distribution of taxa—were crucial to both of these developments. The theory of island biogeography emerged as a joint explanation for two biogeographical gradients: the pattern of increase in species richness on islands with increasing island *size* and with decreasing *isolation* from source areas. MacArthur's fascination with the most striking biogeographical gradient on the planet, the increase in species richness with decreasing *latitude*, began with a series of papers proposing theoretical links between this gradient and community ecology (MacArthur 1965, 1969; Klopfer and MacArthur 1960, 1961) and culminated in pivotal chapters of *Geographical Ecology* that also brought to bear the key concepts of gradients in area and isolation, derived from island biogeography. Although not all of MacArthur's proposals have stood the test of time, these works challenged biogeographers to think theoretically and abstractly about biogeographical gradients.

With the further development of biogeographical databases, rapidly growing computational power, and the development of inexpensive GIS software for personal computers in the 1980s, the stage was set for two key advances in the theory of biogeographical gradients. The emergence of macroecology (Brown and Maurer 1989, Brown 1995, Gaston and Blackburn 2000) as a discipline formalized the intersection, already underway, between broad spatial scales and ecological mechanisms and inferences from large empirical datasets (e.g. Currie and Paquin 1987, Currie 1991). Meanwhile, explicitly stochastic models of species distributions on biogeographical scales (Pielou 1977, Colwell and Winkler 1984, Colwell and Hurtt 1994, Willig and Lyons 1998, Pineda and Caswell 1998, Bokma et al. 2001, Hubbell 2001) stimulated new approaches to simulation modeling on biogeographical scales. In addition to bringing stochastic processes into focus, a key outgrowth of these approaches has been a shift towards modeling biogeographical patterns at the level of species ranges or populations, treating species richness as a function of the overlap of geographical ranges in geographical space, rather than a black-box emergent property.

DOMAIN OF THE THEORY AND MODELS

This chapter aims to review the current role of biogeographical gradients within the realm of ecological theory, recognizing that biogeography has broad and deep connections with the disciplines of evolutionary biology, paleontology, and earth history. The domain of biogeographical gradient theory treated in this chapter encompasses the *characteristics and causes of spatial gradients in the occurrence and co-occurrence of species on geographical scales*. In this context, "geographical scales" refers to spatial scales from the level of ecoregions, biomes, and geographical regions to continents and ocean basins. Chapter 13 (Fox et al.) treats the related, but distinct concept of ecological gradients, which focus on conditions and resources, whether physical, chemical, or biological in nature. The key distinctions lie in the explicit, broad-

scale spatial context of biogeographical gradients, and the restriction of focus, here, to theories and patterns of gradients in species occurrence. At the outset, however, it must be said that it would be impossible to make sense of biogeographical gradients without understanding the role of ecological gradients on broad spatial scales. Because Chapter 10 (Sax and Gaines) treats the important special case of island biogeography, I will focus here on general models of biogeographical gradients.

PROPOSITIONS

- *Proposition 1:* Biogeographical gradients arise from demographic and evolutionary processes acting at the level of populations, including migration, adaptation, speciation, and extinction.
- *Proposition 2:* Biogeographical gradients are composite manifestations of the location and overlap of geographic ranges.
- Proposition 3: Biogeographical gradients have multiple, interacting causes.
- *Proposition 4:* Biogeographical gradients cannot generally be studied experimentally, but statistical analyses and simulations can help reveal their causes.
- *Proposition 5:*Biogeographical gradients characterize the physical world either as spatially explicit patterns or as ordered elements of a *biogeographical mosaic*
- *Proposition 6:* Biogeographical gradients are the consequence of ecological and evolutionary responses of species and lineages to environmental gradients and mosaics.
- *Proposition* 6B = new 7: Biogeographical gradients are shaped by the contingent facts of both earth history and human history.
- *Proposition 8:* Species ranges and the biogeographical gradients that arise from them in physical space can be mapped and modeled on abstract environmental gradients in niche space.
- *Proposition 9:* Because biogeographical gradients are realized in spatially bounded domains, constraints of geometry may affect them, independent of biological and historical influences.

GRADIENTS OF SPECIES RICHNESS

Geographical patterns of species richness have seized the attention of naturalists since the earliest days of European natural history exploration, particularly the latitudinal gradient in

richness (Hillebrand 2004) that characterizes most taxa (e.g. von Humboldt 1807, Wallace 1878). The remarkable richness of tropical biotas, from rainforests to reefs, contrasted with the comparative biotic simplicity of temperate and boreal biotas has inspired a wealth of proposed explanations and theories (e.g., Wallace 1878, Dobzhansky 1950, MacArthur 1965, Pianka 1966, Huston 1994, Rosenzweig 1995), most recently reviewed by Willig et al. (2003). Geographical gradients of increasing richness from dryer to wetter climates, doubtless familiar for millennia to travelers, offer a second repeated pattern in need of explanation (O'Brien 1998, Hawkins et al. 2003a). Changes in richness with elevation on land (Rahbek 1995, 2005), and with depth in the oceans (Pineda and Caswell 1998, Levin et al. 2001), although even more complex and varied than the latitudinal or dry-wet gradients, provide additional examples of repeated biogeographical gradients that have been the focus of intensive study.

Two key concepts in island biogeography theory (see Sax and Gaines Chapter 10) are the decrease in species richness with isolation from biotic source areas and the increase in species richness with island area. The isolation effect also plays a role in continental biotas when topographic or climatic factors isolate regions from sources of colonists (Lomolino et al. 1989) or when distance *per se* plays a role in preventing species from reaching suitable locations. In a more general sense, isolation is an expression of dispersal limitation (Ehrlen and Eriksson 2000, Shurin 2000), which is a fundamental characteristic of living things. The increase of richness with area is equally general and enjoys an enormous literature of it own in under the rubric of species-area relations (SARs) (Rosenzweig 1995, He and Legendre 1996, Scheiner 2003). In this chapter I will touch on the role of area as it is thought to affect biogeographical gradients.

THEORIES AND MODELS OF GEOGRAPHIC GRADIENTS OF SPECIES RICHNESS

The hardening of ecology into a self-consciously experimental science in the last several decades of the twentieth century (e. g. Hairston 1989), with its attendant enthusiasm for careful design and rigorous analysis, raised the bar for what counts as explanation in the discipline. Theories and the models that embody them count too, of course, but they are traditionally viewed as being on a different plane from experiments, and some of the most-admired work combines analytical models with experimental field studies (e.g. Simberloff & Wilson 1969). Unfortunately, the spatial and temporal scales on which physically manipulative experiments are feasible (and the relatively constrained circumstances under which they are ethical) place such experiments out of

reach for many compelling and longstanding questions, including the causes of species richness patterns on geographical scales. Models offer the way forward for the investigation of such questions, and certain kinds of models permit experimentation, as I will explain later.

Statistical models

Statistical models have been widely used to study the correlates, and where possible to suggest the multiple causes (Propositions 3 and 4), of geographical gradients of species richness (e. g., Currie et al. 2004). This approach can be conceived as treating biogeographical gradients as ordered mosaics (Proposition 5) of spatial conditions and represents a clear point of contact between biogeographical gradients and environmental gradients (Fox et al. Chapter 13). For example, a continental-scale map of actual evapotranspiration (AET) might be gridded at a specified resolution (e.g. 2° latitude by 2° longitude cells, or equal-area cells). A scatterplot of the number of bird species (for example) recorded from each cell, plotted against AET for each cell, amounts to ordering the cells (along the abscissa) by AET: an ordered mosaic (e.g. Hawkins et al. 2003b, Figure 14.1). Multivariate environmental correlates have often been modeled as a multidimensional, ordered mosaic, with species richness (at some specified spatial scale) treated as the response variable. From such statistical models emerge linear or more complex functions fitted to the data, from which the richness of individual map cells may deviate positively or negatively.

Because nearby map cells are more likely share similar conditions and similar levels of species richness than are distant map cells, spatial autocorrelation, if not taken into account, can inflate the apparent sample size and Type 1 error. Regardless of spatial autocorrelation in the environment or in richness, however, it is the regression *residuals* that matter: if they show substantial spatial autocorrelation (by Moran's I or other measures), then spatial regression methods must be used to account for any unexplained spatial structure remaining in the regression residuals. On the other hand, if the model residuals show little or no spatial autocorrelation in the best models (that is, the models explain the data very well), then spatial regression methods are unnecessary (Diniz Filho et al. 2003, Rangel et al. 2006).

Using statistical models applied to ordered multivariate mosaics, a substantial body of work has demonstrated highly significant correlations between species richness and environmental variables, particularly (for terrestrial habitats) measures of solar energy, available water, productivity, and topographic heterogeneity (e.g. Wright 1983, Currie and Paquin 1987, Currie 1991, O'Brien 1998, Rahbek and Graves 2001, Jetz and Rahbek 2002, Hawkins et al. 2003b, Storch et al. 2006, Davies et al. 2007). These statistical models do not, in themselves, constitute causal theories, but have inspired mechanistic theories that make additional testable predictions (Huston 1994, Rosenzweig 1995, Willig et al. 2003, Currie et al. 2004). These ideas are discussed and evaluated by Fox et al. (Chapter 13).

At present, however, the widely documented correlation between productivity-related climatic variables and richness still lacks a well-supported, mechanistic explanation (Currie et al. 2004). Topographic heterogeneity, on the other hand, as a regional level promoter of isolation and speciation (Propositions 1 and 6), has behind it a solid body of natural history and phylogeographic evidence (e.g. Graves 1985, Fjeldsa 1994, Rahbek and Graves 2001, Hughes & Eastwood 2006).

Simulation models and experiments

As an alternative to statistical models, spatially explicit models can be designed to predict biogeographical gradients of species richness (in one or more spatial dimensions) by modeling the underlying geographical ranges, or even the populations of individual species (Propositions 1 and 2). Richness emerges from the model as the overlap of ranges as expressed by joint occupancy of map cells. Ecologists and statisticians (Peck 2004, Grimm et al. 2005; Clark & Gelfand 2006) have recently followed philosophers of science (Winsberg 1999, 2001, 2003), physicists, climate modelers, and others in arguing that experimental methods need not be limited to concrete systems but can be legitimately applied to computer simulations carefully constructed to reflect known, underlying processes (Proposition 4).

Experiments on simulated systems, just like experiments on real ones, are constantly in dialogue with theory and abstract models on the design side, and with statistics on the results side. Astrophysicists, for example, were interested in the unusual convective structure of giant red stars (Winsberg 2001), on which manipulative experiments are not feasible. Although internal convective patterns are known and the laws of fluid dynamics that govern convection are well understood, modeling the complex internal turbulence of such a star proved analytically intractable. Instead, a spatially-explicit, discrete-time simulation model was constructed, simplifying where possible (ignoring the internal dynamics of the core and treating it as a simple
heat source). Parameters were varied experimentally to discover which factors most influenced the convective patterns and which parameter values produced simulated results that best matched observed ones.

Biogeographical simulation modeling calls upon the same principles of strategic simplification, experimental exploration of parameters, and statistical analysis of results as this example from astrophysics (a field in which manipulation of the physical objects of study is even more impractical than in biogeography), to assess mechanistic hypotheses about biogeographical patterns (Proposition 3).

The Neutral Model: no niches, with and without boundaries

The most fundamental of biogeographical simulation models are the spatially explicit, individual-based "neutral models" of Hubbell (2001) and Bell (2000), in which speciation, extinction, and migration take place on a grid, but the spatial domain is unbounded, with no niche differentiation—or, viewed another way, all species share an identical niche. Beta diversity—change in species composition with distance—develops, but only because of dispersal limitation, not because of environmental gradients, of which there are none. Alpha diversity—the number of species occurring in a unit area—does not vary spatially, nor does mean population

size or geographic range size, averaged over species (Proposition 1). Realistically considered, a classic neutral model is an appropriate baseline for a reasonably homogeneous ecoregion, but not for explicitly, environmentally heterogeneous regions or gradients (Hubbell 2005). Local communities are imbedded in broader "metacommunities," but the spatial domain is effectively unbounded, and "edge effects" are avoided or intentionally ignored (e.g. Chave and Leigh 2002, Solé et al. 2004, Uriarte et al. 2004).

Imposing a spatial boundary on a neutral model of community structure is a simple step towards added realism as a biogeographical



Figure 14.1. Spatial patterns in species richness in a 50x50 virtual grid system, after 3000 time steps, averaged among 10 replicates, when species have shortdistance migration capacity (From Rangel & Diniz-Filho 2005b).

model (Proposition 9). Implementing Hubbell's (2001) model in a bounded lattice (using the genealogical version of the neutral model of Etienne & Olff [2004]), Rangel & Diniz-Filho (2005b) showed that, with dispersal limitation (reasonably short migration distances in relation to domain size), steady-state species richness is no longer uniform across the domain, as in classic neutral models, but instead peaks in mid-domain: a simple (perhaps the simplest possible) biogeographical gradient in species richness (Figure 14.1). The decline in richness towards the boundary is caused by increasingly asymmetric migration, making species loss more frequent nearer the boundaries after local disappearance following disturbance events (Proposition 1). In Rangel & Diniz-Filho's (2005b) bounded neutral model, the range for each species can be defined the same way that ranges are specified on real-world maps: as minimum polygons surrounding all the occurrences of each species. When patterns from the bounded neutral model are analyzed in this way, the mid-domain peak of richness remains (Propositions 2 and 8).

Range-based models: still no niches

Modeling at the species (range) level, directly, rather than at the individual level, greatly simplifies the prospect of building simulation models that add realism by integrating nichedriven processes and historical contingencies (Proposition 2). The simplest range-based stochastic models of species richness gradients, however, incorporate neither niche differentiation nor history, and model range location on a one-dimensional spatial domain. Colwell and Hurtt (1994) first showed that random placement of one-dimensional "ranges" (line segments) within a one-dimensional "geographical domain" (a bounded line) produces a peak of range overlap (richness) in the middle of the line segment, a pattern that Colwell and Lees (2000) later named the *mid-domain effect* (MDE) (Proposition 9). Willig and Lyons (1998) developed a simple mathematical model for one-dimensional MDE models, for hypothetical ranges, and Pineda and Caswell (1998) introduced the key idea of shuffling empirical ranges at random within a bounded, one-dimensional domain. Lees et al. (1999) developed an analytical method for this *range-shuffling* approach, which they applied to fauna of the Madagascan rainforest biome, strongly implicating MDE as the primary driver of both latitudinal and elevational mid-domain richness peaks (see also Kerr et al. 2006, Lees and Colwell 2007, Currie and Kerr 2007).

Two-dimensional stochastic, range-based models also produce mid-domain richness peaks in a homogenous domain, as initially shown by Bokma et al. (2001) and Jetz and Rahbek

et al. (2001) (Proposition 9). In the latter study, cohesive (continuous) hypothetical ranges were created within a bounded domain (the map of Sub-Saharan Africa), sampling without replacement from the empirical range size frequency distribution (RSFD) for Sub-Saharan African birds. The ranges were simulated by an algorithm inspired by a verbal description by Gotelli and Graves (1996, p. 256), which later came to be known as the *spreading dye algorithm* (Connolly 2005). In this algorithm, *n* hypothetical ranges are randomly placed within the domain, each matching in size (but not in location) one of the *n* empirical ranges for some group of real species endemic to the domain.. For each range, a map cell is chosen randomly, then contiguous cells are added one at a time to the range, until its size equals that of the empirical range it matches.

Whether in one or two dimensions, the mid-domain effect arises from geometric constraints on the location of species ranges within a bounded domain (Proposition 9). The larger a range, the more constrained its location within the domain, such that larger ranges tend to overlap towards the center of the domain. For a unit-line domain, the midpoint of a range of length r ($0 < r \le 1$) is geometrically constrained to be located in a region of length 1 - r in the middle of the domain. For this reason, in a domain with environmental gradients, a key prediction of MDE theory is that the location of smaller ranges on a domain is expected to be influenced more by the environment and less by geometric constraints (Proposition 3), compared with larger ranges (Colwell et al., 2004, 2005).

MDE models were originally intended as null models (Colwell and Hurtt 1994, Willig and Lyons 1998, Lees et al. 1999, Colwell and Lees 2000), from which deviations could be interpreted as caused by environmental or historical gradients (e.g. Connolly et al. 2003). From both a statistical and mechanistic point of view, however, geometric constraints are better treated as explanatory factors, in a multivariate context, on a par with environmental gradients and historical effects (Jetz and Rahbek 2002, Colwell et al. 2004, 2005) (Proposition 3). To this end, the effects of constraints are modeled, with the "niche-shuffling" algorithm of Pineda and Caswell (1998) on one dimensional transects, or with the spreading dye algorithm for two dimensions. This approach has been controversial (Hawkins and Diniz-Filho 2002; Hawkins et al. 2005; Zapata et al. 2003, 2005; Sandel and McKone 2006; Currie and Kerr 2008), in part because of divergent approaches to inference. Some authors (most recently Currie and Kerr 2008), appear committed to a strictly Popperian approach, when it comes to MDE. They view MDE as a "null hypothesis," subject to rejection unless empirical patterns conform within statistical limits to the predictions of an MDE model, while, paradoxically, treating interacting environmental and historical factors as jointly explanatory in a model-selection approach (Proposition 3).

Niche-based models on a bounded environmental gradient

Species interact with their environment in neutral models of community assembly only by undirected drift. As soon as even the simplest environmental gradient and niche differentiation among species is introduced to a model, the rules change. Although classic, demographic models of the evolution of range size on gradients have played a key role in our understanding the interaction of selection and migration on gradients (Kirkpatrick and Barton 1997, Case and Taper 2000), for the study of biographical gradients of species richness, models (to date) have been constructed at the level of species (their niches and geographical ranges), rather than individuals. Ranges are represented by occupied cells on a gridded domain, with range expansion or contraction as a proxy for population dynamics (Propositions 1 and 2).

The gradient model of Rangel & Diniz-Filho (2005a) is perhaps the simplest possible dynamic, niche-based, evolutionary model of biogeographical gradients of species richness. This model sets up a monotonic environmental "suitability" gradient on a bounded, linear domain (in effect, a transect, represented as a row of adjacent cells). The "best" environment is at one end of the domain, where initial species establishment and subsequent speciation is stochastically favored, to a degree controlled by the slope of the linear environmental suitability gradient . Each species is limited in its geographic range (number of occupied cells) by an interaction between the steepness of the environmental gradient and a species-specific, environmental tolerance—its niche breadth (following Kirkpatric & Barton 1997). Niche shifts (adaptation) occur instantaneously at the time of speciation: each new species assumes as its niche optimum the gradient value of its root cell, with a stochastically assigned niche breadth (Propositions 1 and 6). Species do not interact.

The conventional prediction—that the peak of richness will appear at the more "suitable" end of the domain after a period of random speciation and extinction (e.g. Currie 1991, Currie et al. 2004)—is realized only when the environmental gradient is very strong, forcing species to have small ranges, given their environmental tolerances (Figure 14.2a). With weaker gradients,

ranges are larger, given the same tolerances, and a richness peak appears toward the "suitable" end of the gradient. With successively weakened gradients, the richness peak shifts toward the center of the domain (Figure 14.2b-e). As the environmental gradient weakens, ranges become larger (Kirkpatrick and Barton 1997), and the influence of geometric constraints becomes stronger relative to direct effects of the environmental gradient on richness patterns. The mid-domain peak disappears entirely, and the



species richness pattern becomes flat, only at precisely zero gradient strength (Figure 14.2f). Similar results emerge from Connolly's (2005) process-based analytical models.

This simulation model demonstrates the interaction between simple niche structuring and adaptation on gradients and the stochastic effects of traditional Neutral Model processes (speciation, extinction, and implicit range dynamics) within the geometric constraints imposed by a bounded domain (Propositions 1, 2, 3, and 6). Models like this one and its two-dimensional analogues (next section) offer an opportunity to explore a challenging and often contentious issue in quantitative biogeography: the detection of the relative influence of candidate explanatory factors for species richness, which can be investigated rigorously and experimentally in a model system, such as this one, where the causes of pattern are known with certainty (Proposition 4).

Range-based models in a bounded, environmentally heterogeneous domain

Models in this category add realism—and complexity—by replacing the linear gradient of the simplest niche-based gradient models (e.g. Figure 14.2) with a bounded, heterogeneous, environmental mosaic in two dimensions (Proposition 6). Although hypothetical landscapes

would offer more control of environmental patterns, models in this category, to date, have been based on gridded maps of real continents or regions (*environmental maps*), and are thus environmentally multivariate and realistically complex spatially.

Two distinct (and incompatible) approaches have been taken to modeling species distributions, and thus species richness, on environmental maps. In the first, an empirical range size frequency distribution (RSFD) for some particular taxon for the modeled domain is used as the basis for the modeled distributions. This approach has its roots in the work of Pineda and Caswell (1998) and Lees et al. (1999). For example, to investigate the drivers of gradients of bird species richness worldwide, Storch et al. (2006) used the empirical RSFD for all land birds, and a worldwide terrestrial map of actual evapotranspiration (AET). Independently, Rahbek et al. (2007) used the empirical RSFD for South American land birds to model avian species richness for multifactor environmental maps of the continent. In both studies, the list of actual range sizes was used, one by one, to map a hypothetical range of the same area on the map following a modified form of Jetz and Rahbek's (2001) spreading dye algorithm in which stochastic range location is guided by the magnitude of one or more mapped environmental variables. Davies et al. (2007) followed a similar approach. Although in different ways, each of these studies considered MDE as a potential explanatory factor in driving spatial patterns of species richness.

The second approach explicitly models evolutionary processes of speciation, range expansion, range shift with environmental change, and extinction, at the species level. In this approach, not only patterns of species richness, but the range size frequency distribution arises from the biogeographical dynamics of the model itself (Propositions 1 and 6). In this class of models, there is no intended one-to-one correspondence with any particular empirical species, but simulated species are instead viewed as collectively representative of a chosen empirical taxon being modeled. Bokma et al. (2002) developed a predecessor of this approach—a cellular automaton model in a domain without environmental gradients. The earliest model in this class to simulate evolutionary processes on an environmentally textured domain was developed by Brayard et al. (2005), for foraminifera in the Atlantic Ocean. These authors showed that an interaction between sea surface temperature, currents, and MDE—for both the shape of the domain and for temperature—can produce the twin peaks of richness, north and south of the Equator, that characterize the empirical pattern of richness for these organisms (Propositions 3 and 8).

Models in this group have begun to incorporate temporal variation in the environment, both stochastic and periodic (Proposition 7). In the model of Rangel et al. (2007), empirical environmental maps undergo simultaneous, sinusoidal variation in all factors. Tolerance for and adaptation to these shifting environmental factors is modeled directly in multivariate niche space (Hutchinson 1957, 1978; see Chase Chapter 5), whereas range expansion, range fragmentation, speciation, and extinction are modeled as projection from niche space onto geographic space (Pulliam 2000). This powerful reciprocal correspondence (mapping) between the cells of an *n*factor set of environmental maps (in GIS terminology, a map with *n* environmental layers), and the corresponding *n*-dimensional niche space, which has been called *Hutchinson's duality* (Proposition 8) (Colwell and Rangel, in revision)., lies at the heart of dynamic, mechanistic models of biogeographical gradients, and relates them conceptually to ecological gradients (Fox et al. Chap. 13).

In the Rangel et al. (2007) model, for example, "life" originates in a single, randomly chosen starting cell in the geographical space, the gridded environmental map of South America. Based on its n environmental values (layers), this initial map cell corresponds to a single point in the corresponding *n*dimensional environmental niche space, in which the point is declared the niche center for this founder species. A niche breadth for each factor is assigned independently and stochastically for each environmental axis in niche space. The niche hypervolume thus defined is projected back onto map space, defining the geographical range of the seed species as all cells on the map that are *contiguous* with the



the best-fitting model. (C) Relationship and OLS (ordinary least squares) fit between the patterns in maps (A) and (B) (Rangel et al. 2007).

founder species' seed cell, and that lie within its niche in niche space (Proposition 8). As the environmental factors fluctuate, the range of the founder species becomes fragmented, traversed by areas of the map that no longer correspond to areas within its niche (Proposition 8). Smaller fragments face stochastic extinction; larger ones become daughter species. The model allows each surviving fragment to adapt to its new conditions and expand its new niche in niche space around its own niche center, to a degree controlled by a "niche conservatism" parameters in the model. The dynamic processes of range fragmentation, extinction, speciation, and constrained adaptation continue until some specified number of species ranges populate the geographical space (Propositions 1 and 6). The model thus produces a phylogeny of niches in niche space, mapped by Hutchinson's duality (Proposition 8) into geographical space, that invites further

exploration. With the empirical species richness map for South American birds as a criterion (Figure 14.3), Rangel et al. (2007) treated their exploration of model parameter space as a multi-factorial, muti-level experiment (Proposition 4), concluding that a relatively low extinction rate, a substantial level of niche conservatism (Wiens and Donoghue 2004), and an equatorial latitude for the founder species' seed cell are key conditions for a good fit between the observed and modeled geographical patterns of species richness of birds (Proposition 3).

BIOEOGRAPHICAL GRADIENTS OF RANGE SIZE

Biogeographical gradients in range size and species richness are difficult to disentangle, and the quest to understand their interactions has motivated the history of quantitative biogeography for decades. Moreover, from a theoretical viewpoint, it makes little sense to consider static biogeographical patterns of range size without considering the roles of adaptation, speciation, and extinction in relation to range size (Propositions 1 and 6), as discussed in the previous section. These are complex topics, encompassing a large literature. I offer here only an outline of the theoretical underpinnings.

In *Aerografía*, his ground-breaking, but long-underappreciated monograph, Eduardo Rapoport (1975, English translation in 1982), reported that latitudinal range size tends to be smaller for tropical subspecies than for temperate subspecies, within species of New World mammals and three orders of birds in Asia. As applied to species, rather than subspecies (the difference not always being obvious, in any case), Stevens' (1989) canonized this pattern as *Rapoport's rule*. Stevens presented several examples of the same qualitative pattern, all for north temperate groups (latitude 25°N and beyond), which he documented by regressing mean range size against latitude for latitudinal bands (Proposition 2).

Two decades after Stevens' (1989) paper and more than three after Rapoport's (1975) farsighted explorations of geographical ranges, we can now look back and appreciate their profound effect on the theory of biogeographical gradients, primarily by making us all try to think more clearly about how the geography of ranges produces the geography of richness (Proposition 2). On the methodological side, the seemingly straightforward concept that Stevens put forward quickly ran into trouble, coming as it did at a time when we ecologists were coming to our senses regarding spatial autocorrelation (Pagel et al. 1991, Legendre 1993), pseudoreplication (Hurlbert 1984), and phylogenetic non-independence (Harvey and Pagel 1991). Because Stevens' approach counts the same ranges repeatedly (in proportion to their size), the variates for Stevens' regressions are not statistically independent, and degrees of freedom are thus inflated. Colwell and Hurtt (1994), using an approach introduced by Graves (1985), suggested, instead, a "midpoint plot," in which each species is plotted once, in a scattergram of range size vs. latitudinal midpoint (Colwell and Hurtt [1994], Figure 9). The midpoint plot made clear that range size and range midpoint are inherently non-independent: a range with a high-latitude midpoint cannot be as large a range with a midpoint nearer the equator (Proposition 9). Exploration of this geometric constraint, in the context of Rapoport's rule, was the key to the discovery of the mid-domain effect, independently, by both Colwell and Hurtt (1994) and by Lyons and Willig (1997).

Rohde et al. (1993) compared Steven's method (range size mean for species occurring in each latitudinal band) to what they called the "midpoint method" (range size mean for only those species whose midpoint occurs in each latitudinal band). They not only found no support for Rapoport's rule for marine teleost fishes, by either method, but showed that species with tropical range midpoints have broader ranges than species with midpoints at temperate latitudes. Although these two methods measure different things (Colwell and Hurtt 1994, Connolly in press), they often produce qualitatively concordant results for empirical data.

While avoiding statistical non-independence and spatial autocorrelation due to repeated contributions from larger ranges in the Stevens method, the midpoint method does nothing to deal with phylogenetic non-independence. In an assessment of Rapoport's rule for New World endemic land-birds, Blackburn and Gaston (1996) attempted to account for this problem by using phylogenetically independent contrasts. They compared results of this method with three others:: Stevens' method, Colwell and Hurtt's (1994) midpoint scatterplot method (which Blackburn and Gaston named the "across species method"), and Rohde et al.'s (1993) midpoint (band mean) method. Independent of method, they reported that range size indeed reaches a minimum in the tropics, but at about 12°N latitude (Nicaragua), not at the equator, a pattern that the authors attribute to biogeographical history (Proposition 7), rather than the mechanisms that Stevens conjectured to be driving Rapoport's rule (discussed below).

However, the coarse data quality (published range maps, rather than primary data) and the very large spatial grain of the analysis (10° cells) used by Blackburn and Gaston (1996) and in numerous other studies blinds the analysis to the spatial pattern of small-ranged species. For example, nearly 70% of the 241 species of South American hummingbirds, the majority Andean, have ranges smaller than a single 10° quadrat (Rahbek and Graves 2000). This source of error is not a random one, but would tend to bias the against extending the pattern of decreasing range size (Rapoport's rule) into the equatorial tropics. On the other hand, primary survey and collections data for most taxa, unless corrected for it, are likely to display the opposite bias (underestimation of range size in the tropics), because of under sampling in rich communities (Colwell and Hurtt 1994).

Colwell and Hurtt (1994) had shown that, depending upon the range size frequency distribution and random placement algorithm, a reverse Rapoport effect (larger ranges in the tropics, smaller ranges at higher latitudes) can appear simply as a result of geometric constraints (Proposition 9), using Steven's method of analysis. Lyons and Willig (1997) took the non-independence of midpoint and range into full account through simulation, and concluded, for midpoint plots, that New World bats and marsupials both support Rapoport's rule.

The conflicting results from these studies are typical of what was already quite a substantial literature on the subject when reviewed a decade ago (Gaston et al. 1998). Published empirical support for the "rule" turns out to be variable, not only taxonomically but also geographically, with land areas north of the Tropic of Cancer offering the strongest evidence for it. Existing studies covering tropical latitudes offer considerably less support, but most of these suffer from data quality and methodological issues, suggesting that the "epitaph" famously called for by Gaston et al. (1998) may in the long run proved to be a case of premature burial (Poe 1844). On the other, the conceptual side of Stevens' (1989) seminal paper appears to have survived the challenges and controversies surrounding the prevalence and detection of the empirical pattern.

Stevens went beyond describing Rapoport's rule as an empirical pattern, to conjecture that the pattern was a key to understanding latitudinal (Stevens 1989), and later, elevational (Stevens 1992) and depth (Stevens 1996) gradients of species richness. His idea was founded on the landmark "mountain passes" paper of Janzen (1967), who speculated that tropical species, because they have evolved in climates (particularly temperature regimes) that vary little seasonally compared with seasonal patterns at higher latitudes, would tend to have narrower climatic tolerances than species at higher latitudes (Proposition 6). Stevens proposed that these narrow tolerances (particularly temperature, but also precipitation) would restrict tropical species to smaller geographical ranges than the wider climatic tolerances expected at higher latitudes (Proposition 8)—a mechanism that he supposed to explain Rapoport's findings for vertebrate subspecies. If geographical ranges are small, Stevens argued, then demographic sink areas, lying outside the range limits for positive fitness, would overlap broadly among many species, yielding areas of high local species density. This hypothesis, which is logically quite separate from the range-size pattern described by Rapoport's rule, has come to be known as the "Rapoport rescue

effect" (Stevens 1992), after Brown and Kodric-Brown's (1977) idea that demographic sink populations are continually "rescued" from extinction by immigration (Proposition 1).

In a reverse approach, Taylor and Gaines (1999) used stochastic range simulations on a spherical domain to force a "classic" Rapoport effect, with or without demographic sink perimeters on species ranges. They found that the resulting pattern of species richness was opposite the empirical one—the poles were richer than the equator—calling into question Stevens' (1989, 1992) conjecture that the latitudinal gradient in species richness might be a consequence of Rapoport's rule.

For the latitudinal gradient of species richness, Gaston and Chown (1999) pointed out that a fundamental empirical problem undermines Stevens' (1989) appealing line of reasoning. Steven's argument implicitly assumes



Figure 14.4. (A) Global map of annuallyaveraged, near-surface air temperature from 1961-1990. (B) Variation in surface temperature as a function of latitude, after removing the effects of varying surface elevation, based on data from (A). Gray regions show first and second standard deviations. Copyright Robert A. Rohde, Global Warming Art, reproduced with permission

that the latitudinal gradient in mean annual temperature, familiar to anyone who has gone south to Mexico or Spain, or north to Queensland, Mombasa, or Rio for a winter holiday, continues all the way to the Equator. In fact, as pointed out by Terborgh (1973) and discussed by Rosenzweig (1995), Gaston and Chown (1999), and Colwell et al. (2008), the latitudinal gradient in mean annual temperature, which rises almost linearly from the poles and to the Tropics of Cancer and Capricorn for lowland continental stations, levels off to a broad plateau within the tropics (Figure 14.4). For this reason, the difference in the scope of thermal tolerance between tropical and temperate species anticipated under Janzen's (1967) hypothesis would not be expected to produce a latitudinal Rapoport pattern in the first place, at least within the tropics, in the absence of other limitations to species ranges, such as topography or simply dispersal limitation. Indeed, in principle, a tropical species with a narrow temperature tolerance might well have a broader latitudinal range than a temperate species with a larger temperature tolerance, given the broad thermal plateau between 20° S and 20°N (Terborgh 1973). The key to understanding this issue is an accurate mapping between niche space (thermal niche breadth, in this case) and environmental map space (Proposition 8). At the very least, based on mean annual temperatures, latitudinal ranges of species restricted to the tropics cannot be assumed to be correlated in any simple way with their temperature tolerances. On the other hand, Steven's reasoning makes more sense for temperate and boreal species, along the nearly-linear portion of the latitudinal temperature gradient. Perhaps it is no coincidence that many of the best cases for Rapoport's rule (Gaston et al. 1998), including every one of Stephens' (1989) examples, covered latitudes beyond 25°N, but a definite answer requires more and better data and analyses that account for geometric constraints, spatial autocorrelation, and phylogenetic non-independence. In contrast with the latitudinal gradient in temperature (for a given elevation), which is mostly a temperate and arctic pattern, the lapse rate (decline in temperature with elevation) differs little with latitude, on a mean annual basis (Colwell et al. 2008). Whether you walk up a tropical mountain or temperate one, the temperature declines roughly the same amount, about 5 $^{\circ}$ C for every 1 km of elevation. However, because seasonal temperature variation at all elevations is greater at temperate latitudes than within the tropics, Janzen (1967) predicted that elevational ranges of tropical species should be considerably narrower than elevational ranges for related temperate species. Surprisingly, it was not until 40 years later, that McCain (2009) carried out a meta-analysis of datasets published for other reasons, and of varying quality and completeness (under sampling in

rich, tropical biotas remains a hazard). Even in the most stringently restricted subset of these data, however, the results of this analysis support Janzen's prediction for vertebrates. Assuming this latitudinal gradient in elevational range sizes is correct, we may expect that the same topographical gradient would yield smaller geographical (mapped) ranges for tropical species than for temperate species (Proposition 8). This expectation puts a different spin on Rapoport's rule, if it is to include tropical species: the rule would be expected to apply only to the degree that elevational, rather than latitudinal, temperature gradients are driving latitudinal gradients of range size. Clearly, the degree to which this conjecture is correct depends upon geographical patterns of topography, not on climate alone (Rahbek and Graves 2001).

What about Stephens' (1992) application of Rapoport's rule and the Rapoport rescue hypothesis to elevational gradients? In the supposed parallel between latitude and elevation upon which Stephens based his argument, the tropical lowlands were assumed to represent the highest richness and smallest mean range size (both latitudinal and elevational) on Earth, with parallel declines in richness and increases in range size with both latitude and elevation. Under this model, elevational gradients at temperate and boreal latitudes would also show a monotonic decrease in richness and a monotonic increase in elevational range size, but scaled to a higher mean range size because of greater seasonality.

There are at least four serious problems with any parallel application of Rapoport's rule to latitudinal and elevational gradients. First, as discussed above, based solely on the evolution of temperature tolerances, there is no reason to expect (and so far, little evidence to show), that latitudinal ranges are generally smaller in the tropics than at higher latitudes, at a given elevation and for similar topography. Second, it is unclear whether range size increases routinely with elevation, particularly in the tropics, although a thorough meta-analysis of worldwide datasets is needed. The evidence Stevens (1992) presents for increasing range size with elevation in the tropics is almost certainly biased by under sampling, which creates a spurious negative correlation between richness and range size (Colwell and Hurtt 1994). Despite this almost inevitable sampling bias, recent quantitative sampling for some 2000 species of plants and insects on a tropical elevational gradient in Costa Rica revealed no conspicuous increase in elevational range size with elevation (Colwell et al. 2008). The third problem, discussed earlier, is that species richness more often peaks at mid-elevations than in the lowlands, regardless of latitude (Colwell and Hurtt 1994; Rahbek 1995, 1997, 2005; Nogués-Bravo et al. 2008). The

fourth issue is that the absolute spatial scale much greater for the same amount of environmental change on a latitudinal gradient than on an elevational gradient, with important demographic and genetic consequences (Rahbek 2005). For temperature, the elevational gradient (per km elevation) at subtropical to high temperate latitudes is nearly 1000 times greater than the latitudinal gradient (per km poleward). (On the ground, the temperature gradient is 100 times greater, even on a 1% slope.) At tropical latitudes, and there is no latitudinal gradient in temperature, so the contrast is even more extreme (Colwell et al. 2008).

In summary, biogeographical gradients in range size very well may be driven, as Stevens' (1989) proposed, by adaptive tolerances to seasonal fluctuations, but the expected translation from thermal tolerance limits in niche space to geographical ranges in map space (Hutchinson's duality (Proposition 8)) requires an accurate environmental map. Steven's idea that tropical richness could be the product of wide "sink" margins around small geographical ranges is not in accord with the wide geographical scope of tropical climates, at any constant elevation. On the other hand, on steep elevational gradients, local richness may well be enhanced by source-sink dynamics (mass effect [Schmida and Wilson 1985] from physically nearby source areas at lower and higher elevations), particularly in the tropics (Rahbek 1997, Kessler 2000, Grytnes 2003).

CONCLUSION

As delimited in this chapter, the theory of biogeographical gradients comprises the patterns and causes of geographic variation in the size and location species ranges and their overlap, which we express as species richness. I have focused on concepts and models, attempting to work from the simpler to the more complex. The underlying demographic, ecological, and evolutionary processes that ultimately determine all biogeographical patterns in nature have played key roles in this exploration, particularly in translating between the evolution of environmental tolerances, which is best modeled in niche space, and the realization of distributions on the planet, as expressed in geographical space (Proposition 8). Although the more sophisticated models have begun to incorporate the role of deep time in biogeography, the role of earth history in shaping biogeographical gradients—through plate tectonics, mountain-building, changing connections between land masses and seas, and climate history—represents an important modeling frontier for theoretical biogeography. I have not attempted to discuss the role of the rapidly growing body of knowledge of phylogenetic history and phylogeographical studies in biogeography, but these

data also promise to enrich modeling and inform theory in the study of biogeographical gradients. Finally, biogeographical models can neither hope to be meaningful, nor can they be rigorously assessed without accurate, carefully compiled data on the distributions of organisms. There is no substitute for good data.

Literature Cited

- Bell, G. 2000. The distribution of abundance in neutral communities. The American Naturalist 155:606-617.
- Blackburn, T., and K. Gaston. 1996. Spatial patterns in the geographic range sizes of bird species in the New World. Philosophical Transactions of the Royal Society B: Biological Sciences 351:897-912.
- Bokma, F., J. Bokma, and M. Mönkkönen. 2001. Random processes and geographic species richness patterns: why so few species in the north. Ecography 24:43-49.
- Brayard, A., G. Escarguel, and H. Bucher. 2005. Latitudinal gradient of taxonomic richness: combined outcome of temperature and geographic mid-domains effects? Journal of Zoological Systematics & Evolutionary Research 43:178-188.
- Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago.
- Brown, J. H., and A. Kodric-Brown. 1977. Turnover rates in insular biogeography: effects of immigration on extinction. Ecology 58:445-449.
- Brown, J. H., and B. A. Maurer. 1989. Macroecology: The division of food and space among species on continents. Science (Washington, D. C.) 243:1145-1150.
- Case, T. J., and M. L. Taper. 2000. Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. American Naturalist 155:583-605.
- Chave, J., and E. G. Leigh. 2002. A spatially explicit neutral model of b-diversity in tropical forests. Theoretical Population Biology 62:153-168.
- Clark, J. S., and A. E. Gelfand. 2006. A future for models and data in environmental science. Trends in Ecology & Evolution 21:375-380.
- Colwell, R. K., G. Brehm, C. Cardelús, A. C. Gilman, and J. T. Longino. 2008. Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. Science 322:258-261.
- Colwell, R. K., and G. C. Hurtt. 1994. Nonbiological gradients in species richness and a spurious Rapoport effect. American Naturalist 144:570-595.
- Colwell, R. K., and D. C. Lees. 2000. The mid-domain effect: geometric constraints on the geography of species richness. Trends in Ecology and Evolution 15:70-76.
- Colwell, R. K., C. Rahbek, and N. Gotelli. 2004. The mid-domain effect and species richness patterns: what have we learned so far? American Naturalist 163:E1-E23.

- Colwell, R. K., C. Rahbek, and N. Gotelli. 2005. The mid-domain effect: there's a baby in the bathwater. American Naturalist 166:E149–E154.
- Colwell, R. K., and T. F. Rangel. In revision. Hutchinson's duality: the once and future niche. PNAS 00:000-000.
- Colwell, R. K., and D. W. Winkler. 1984. A null model for null models in biogeography. Pages 344-359 *in* D. R. Strong, Jr., D. Simberloff, L. G. Abele, and A. B. Thistle, editors. Ecological communities: Conceptual issues and the evidence. Princeton University Press, Princeton, N. J.
- Connolly, S. R., D. R. Bellwood, and T. P. Hughes. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. Ecology 84:2178-2190.
- Connolly, S. R. 2005. Process-based models of species distributions and the mid-domain effect. American Naturalist 166:1-11.
- Connolly, S. R. 2009 (in press). Macroecological theory and the analysis of species richness gradients. Pages 000-000 *in* J. D. Witman and K. Roy, editors. Marine macroecology. University of Chicago Press, Chicago.
- Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. American Naturalist 137:27-49.
- Currie, D. J., and J. T. Kerr. 2007. Testing, as opposed to supporting, the Mid-domain Hypothesis: a response to Lees and Colwell (2007). Ecology Letters 10:E9-E10.
- Currie, D. J., and J. T. Kerr. 2008. Tests of the mid-domain hypothesis: a review of the evidence. Ecological Monographs 78:3-18.
- Currie, D., and V. Paquin. 1987. Large-scale biogeographical patterns of species richness of trees. Nature 329:326-327.
- Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J. F. Guegan, B. A. Hawkins, D. M. Kaufmann et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters 12:1121-1134.
- Davies, R., C. Orme, D. Storch, V. Olson, G. Thomas, S. Ross, T. Ding, P. Rasmussen, P. Bennett, and I. Owens. 2007. Topography, energy and the global distribution of bird species richness. Proceedings of the Royal Society B: Biological Sciences 274:1189-1197.
- Diniz-Filho, J. A. F., B. L.M., and H. B.A. 2003. Spatial autocorrelation and red herrings in geographical ecology. Global Ecology and Biogeography 12:53-64.

Dobzhansky, T. 1950. Evolution in the tropics. American Scientist 38:209-221.

- Ehrlen, J., and O. Eriksson. 2000. Dispersal limitation and patch occupancy in forest herbs. Ecology 81:1667-1674.
- Etienne, R. S. and H. Olff. 2004. A novel genealogical approach to neutral biodiversity theory. Ecology Letters 7:170-175.
- Fjeldså, J. 1994. Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. Biodiversity and Conservation 3:207-226.
- Gaston, K. J., and T. M. Blackburn. 2000. Pattern and process in macroecology. Blackwell Science, Oxford.
- Gaston, K. J., T. M. Blackburn, and J. I. Spicer. 1998. Rapoport's rule: time for an epitaph? Trends in Ecology and Evolution 13:70-74.
- Gaston, K., and S. Chown. 1999. Why Rapoport's rule does not generalise. Oikos 84:309.
- Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution Press, Washington, D. C.
- Graves, G. R. 1985. Elevational correlates of speciation and intraspecific geographic variation in plumage in Andean forest birds. Auk 102:556-579.
- Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H-H. Thulke, J.Weiner, T. Wiegand, and D. L. DeAngelis. 2005. Pattern-Oriented Modeling of Agent-BasedComplex systems: Lessons from Ecology. Science 310:987-991.
- Grytnes, J. 2003. Ecological interpretations of the mid-domain effect. Ecology Letters 6:883-888.
- Hairston, N. 1989. Ecological experiments: purpose, design, and execution. Cambridge University Press.
- Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford.
- Hawkins, B. A., and J. A. F. Diniz-Filho. 2002. The mid-domain effect cannot explain the diversity gradient of Nearctic birds. Global Ecology and Biogeography 11:419-426.
- Hawkins, B. A., J. A. F. Diniz-Filho, and A. E. Weis. 2005. The mid-domain effect and diversity gradients: is there anything to learn? American Naturalist 166:E140-E143.

- Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J.-F. Guégan, D. M. Kaufmann, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003a.
 Energy, water, and broad-scale geographic patterns of species richness. Ecology 84:3105–3117.
- Hawkins, B. A., E. E. Porter, and J. A. F. Diniz-Filho. 2003b. Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. Ecology 84:1608-1623.
- He, F., and P. Legendre. 1996. On species-area relations. American Naturalist 148:719.
- Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. American Naturalist 163:192-211.
- Hubbell, S. P. 2001. The unified theory of biodiversity and biogeography. Princeton University Press, Princeton, N. J.
- Hubbell, S. P. 2005. Neutral theory in community ecology and the hypothesis of functional equivalence. Functional Ecology 19:166-172.
- Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. PNAS 103: 10334-10339.
- Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. 54:187-211.
- Huston, M. 1994. Biological diversity: the coexistence of species on a changing landscape. Cambridge, Cambridge University Press.
- Hutchinson, G. 1957, Concluding remarks Cold Spring Harbor Symposia on Quantitative Biology 22:415-427.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? The American Naturalist 93:145.
- Hutchinson, G. E. 1978. An introduction to population biology. Yale University Press, New Haven.
- Janzen, D. H. 1967. Why mountain passes are higher in the tropics. American Naturalist 101:233-249.
- Jetz, W., and C. Rahbek. 2001. Geometric constraints explain much of the species richness pattern in African birds. PNAS 98:5661-5666.
- Jetz, W., and C. Rahbek. 2002. Geographic range size and determinants avian species richness. Science 279:1548–1551.

- Kerr, J. T., M. Perring, and D. J. Currie. 2006. The missing Madagascan mid-domain effect. Ecology Letters 9:149–159.
- Kessler, M. 2000. Upslope-directed mass effect in palms along an Andean elevational gradient: a cause for high diversity at mid-elevations? Biotropica 32:756-759.
- Kirkpatrick, M. and N. H. Barton 1997. Evolution of a species' range. American Naturalist 150:1-23.
- Klopfer, P., and R. H. MacArthur. 1960. Niche size and faunal diversity. American Naturalist 94:293-300.
- Klopfer, P., and R. H. MacArthur. 1961. On the causes of tropical species diversity: Niche overlap. American Naturalist 95:223-226.
- Lees, D. C., and R. K. Colwell. 2007. A strong Madagascan rainforest MDE and no equatorward increase in species richness: Re-analysis of 'The missing Madagascan mid-domain effect', by Kerr J.T., Perring M. & Currie D.J (*Ecology Letters* 9:149-159, 2006). Ecology Letters 10:E4-E8.
- Lees, D. C., C. Kremen, and L. Andriamampianina. 1999. A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. Biol. J. Linn. Soc. 67:529-584.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659-1673.
- Levin, L. A., R. J. Etter, M. A. Rex, A. J. Gooday, C. R. Smith, J. Pineda, C. T. Stuart, R. R. Hessler, and D. Pawson. 2001. Environmental influences on regional deep-sea species diversity. Annual Review of Ecology and Systematics 32:51-93.
- Lomolino, M. V., J. H. Brown, and R. Davis. 1989. Island biogeography of montane forest mammals in the American Southwest. Ecology 70:180-194.
- Lyons, S. K., and M. R. Willig. 1997. Latitudinal patterns of range size: methodological concerns and empirical evaluations for New World bats and marsupials. Oikos 79:568-580.
- McCain, C. M. 2009. Vertebrate range sizes indicate that mountains may be higher in the tropics. Ecology Letters 12:550-560.
- MacArthur, R. H. 1965. Patterns of species diversity. Biological Review 40:510-533.
- MacArthur, R. H. 1969. Patterns of communities in the tropics. Biological Journal of the Linnean Society 1:19-30.
- MacArthur, R. H. 1972. Geographical ecology: patterns in the distribution of species. Harper and

Row, New York.

- MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, New Jersey.
- Nogués-Bravo, D., M. Araújo, T. Romdal, and C. Rahbek. 2008. Scale effects and human impact on the elevational species richness gradients. Nature 453:216-219.
- O'Brien, E. M. 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. Journal of Biogeography 25:379-398.
- Pagel, M. D., R. M. May, and A. R. Collie. 1991. Ecological aspects of the geographical distribution and diversity of mammalian species. American Naturalist 137:791-815.
- Peck, S. L. 2004. Simulation as experiment: a philosophical reassessment for biological modeling. Trends in Ecology and Evolution 19:530-534.
- Pianka, E. R. 1966. Latitudinal gradients in species diversity: a review of concepts. American Naturalist 100:33-46.
- Pielou, E. C. 1977. The latitudinal spans of seaweed species and their patterns of overlap. Journal of Biogeography 4:299-311.
- Pineda, J., and H. Caswell. 1998. Bathymetric species-diversity patterns and boundary constraints on vertical range distributions. Deep-Sea Research II 45:83-101.
- Poe, E. A. 1844. "The premature burial". The Philadelphia Dollar Newspaper, Philadelphia.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters 3:349-361.
- Rahbek, C. 1995. The elevational gradient of species richness: a uniform pattern? Ecography 19:200-205.
- Rahbek, C. 1997. The relationship between area, elevation and regional species richness in Neotropical birds. American Naturalist 149:875-902.
- Rahbek, C. 2005. The role of spatial scale in the perception of large-scale species-richness patterns. Ecology Letters 8:224-239.
- Rahbek, C., N. Gotelli, R. K. Colwell, G. L. Entsminger, T. F. L. V. B. Rangel, and G. R. Graves. 2007. Predicting continental-scale patterns of bird species richness with spatially explicit models. Proceedings of the Royal Society of London Series B 274:165-174.

- Rahbek, C. and G. R. Graves. 2000. Detection of macro-ecological patterns in South American hummingbirds is affected by spatial scale. Proceedings of the Royal Society of London series B-Biological Sciences 267:2259-2265.
- Rahbek, C. and G. R. Graves. 2001. Multiscale assessment of patterns of avian species richness. PNAS 98:4534–4539.
- Rangel, T. F. L. V. B., and J. A. F. Diniz-Filho. 2005a. An evolutionary tolerance model explaining spatial patterns in species richness under environmental gradients and geometric constraints. Ecography 28:253-263.
- Rangel, T. F. L. V. B., and J. A. F. Diniz-Filho. 2005b. Neutral community dynamics, the middomain effect and spatial patterns in species richness. Ecology Letters 8:783-790.
- Rangel, T. F. L. V. B., J. A. F. Diniz-Filho, and L. M. Bini. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Global Ecology & Biogeography 15:321-327.
- Rangel, T. F. L. V. B., J. A. F. Diniz-Filho, and R. K. Colwell. 2007. Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. American Naturalist 170:602-616.
- Rapoport, E. H. 1975. Areografía: estrategias geográficas de las especies. Fondo de Cultura Económica, Mexico City.
- Rapoport, E. H. 1982. Areography. Geographical strategies of species, 1st English ed. edition. Pergamon Press, Oxford.
- Rohde, K., M. Heap, and D. Heap. 1993. Rapoport's rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. American Naturalist 142:1-16.
- Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.
- Sandel, B. S., and M. J. McKone. 2006. Reconsidering null models of diversity: Do geometric constraints on species ranges necessarily cause a mid-domain effect? Diversity and Distributions 12:467–474.
- Scheiner, S. M. 2003. Six types of species-area curves. Global Ecology & Biogeography 12:441-447.
- Schmida, A., and M. V. Wilson. 1985. Biological determinants of species diversity. Journal of Biogeography 12:1-20.

- Shurin, J. 2000. Dispersal limitation, invasion resistance, and the structure of pond zooplankton communities. Ecology 81:3074-3086.
- Simberloff, D., and E. Wilson. 1969. Experimental zoogeography of islands: the colonization of empty Islands. Ecology 50:278-296.
- Solé, R. V., D. Alonso, and J. Saldaña. 2004. Habitat fragmentation and biodiversity collapse in neutral communities. Ecological Complexity 1:65-75.
- Stevens, G. C. 1989. The latitudinal gradient in geographical range: how so many species coexist in the tropics. American Naturalist 133:240-256.
- Stevens, G. C. 1992. The elevational gradient in altitudinal range: An extension of Rapoport's latitudinal rule to altitude. American Naturalist 140:893-911.
- Stevens, G. C. 1996. Extending Rapoport's rule to Pacific marine fishes. Journal of Biogeography 23:149-154.
- Storch, D., R. G. Davies, S. Zajicek, C. D. L. Orme, V. Olson, G. H. Thomas, T. S. Ding, P. C. Rasmussen, R. S. Ridgely, P. M. Bennett, T. M. Blackburn, I. P. F. Owens, and K. J. Gaston. 2006. Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. Ecology Letters 9:1308-1320.
- Taylor, P., and S. Gaines. 1999. Can Rapoport's rule be rescued? Modeling causes of the latitudinal gradient in species richness. Ecology 80:2474-2482.
- Terborgh, J. 1973. On the notion of favorableness in plant ecology. American Naturalist 107:481-501.
- Uriarte, M., R. Condit, C. D. Canham, and S. P. Hubbell. 2004. A spatially explicit model of sampling growth in a tropical forest: does the identity of the neighbours matther? Journal of Ecology 92:348-360.
- von Humboldt, A. 1807. Views of nature: or contemplations on the sublime phenomena of creation; with scientific illustrations, 3rd ed., Otté. E.C. and Bohn, H.B., translators (1850), published by Henry G. Bohn,
- Wallace, A. R. 1878. Tropical nature and other essays. Macmillan & Co., London.
- Wiens, J., and M. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends in Ecology & Evolution 19:639-644.
- Willig, M. R., and S. K. Lyons. 1998. An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. Oikos 81:93-98.

- Willig, M. R., D. M. Kaufmann, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. Annu. Rev. Ecol. Syst. 34:273-309.
- Winsberg, E. 1999. Sanctioning models: the epistemology of a simulation. Science in Context 12:247-260.
- Winsberg, E. 2001. Simulations, models, and theories: complex physical systems and their representations. Philosophy of Science 68 (Proceedings):S442-S454.
- Winsberg, E. 2003. Simulated experiments: methodology for a virtual world. Philosophy of Science 70:105-125.
- Wright, D. H. 1983. Species-energy theory: an extension of species-area theory. Oikos 41:496-506.
- Zapata, F. A., K. J. Gaston, and S. L. Chown. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. J. Animal Ecology 72:677-690.
- Zapata, F. A., K. J. Gaston, and S. L. Chown. 2005. The mid-domain effect revisited. American Naturalist 166:E144–E148.

Chapter 15: The State of Theory in Ecology Michael R. Willig and Samuel M. Scheiner

"There is a need to subject current theory to stringent empirical test, but ecology can never have too much theory." Robert H. MacArthur

An article in *Wired Magazine* proclaimed in its title "The End of Theory: the Data Deluge Makes the Scientific Method Obsolete" (Anderson 2008). The basic premise of that essay was that evolving computational capabilities will allow large and heterogeneous data sets to be mined efficiently and effectively. The result would be the production of pattern without the need of hypothesis formation and testing, so much so that correlation would supersede causation. Science would "advance without coherent models, unified theories, or really any mechanistic explanation at all". Although advanced cyberinfrastructure will revolutionize much of the scientific enterprise as it relates to data collection and visualization, the overarching thesis of the article invites replies on many levels. Rather than do so here, we use that statement as a springboard from which to emphasize the unsophisticated view held by many that there is but one scientific method and that the accuracy and utility of models are the essential characteristics of theory. Indeed, perusal of the preceding chapters in this book or the influential tome by Pickett et al. (2007) suggests that Anderson's (2008) conceptualization of theory is flawed from many perspectives. Mark Twain, upon reading his obituary in a New York journal, is quoted as having replied, "the reports of my death are greatly exaggerated." So too, this is our response to Anderson's contention. Theory in general, and theory in ecology in particular, are alive and flourishing, providing much impetus to deepen and broader our understanding of the natural world. Our goal in this chapter is to briefly summarize where the discipline now stands with regard to that understanding.

THE HIERARCHY OF ECOLOGY

The domain of a theory defines its central focus. In the case of a general theory, that domain also circumscribes the scientific discipline that addresses that domain. In Chapter 1, we defined the domain of the theory of ecology as the spatial and temporal patterns of the distribution and abundance of organisms, including their causes and consequences. Nonetheless, this definition

requires additional exposition to understand the nature of the patterns and processes under consideration (Kolasa and Pickett 1989). Specifically, the brief definition of ecology's domain does not address the nature of the interactions that define the levels of the ecological hierarchy, and how that hierarchy fits within the relationships and interactions that define the rest of the domain of biology (Figure 15.1).

Our hierarchical perspective makes clear that ecological theory is directed at understanding biological entities at or above the level of individuals. The birth, death, growth and movement of individual organisms give rise to the complex spatial and temporal tapestry of life that is the focus of ecological studies, and these basic attributes arise from the dynamics involved in the acquisition of energy and nutrients from the environment. Most of the chapters in this book focus on the middle of that hierarchy: populations (Hastings Chapter 6; Holt Chapter 7), single communities (Chase Chapter 5; Pickett et al. Chapter 9), or collections of communities (Leibold Chapter 8; Sax and Gaines Chapter 10; Fox et al. Chapter 13; Colwell Chapter 14).

A few chapters examine theory associated with the ends of the hierarchy. At the level of individuals, the domain of the theory of ecology intersects with the domain of the theory of organisms (Scheiner submitted; Zamer and Scheiner in prep.), resulting in such disciplines as physiological ecology. The theories that define those disciplines are examples of how the domains of a constitutive theory can overlap the domains of more general theories. In this book, the individual-level perspective is represented by foraging theory (Sih Chapter 4), which is representative of the broader domain of behavioral ecology.

At the other end of the hierarchy, the theory of ecology overlaps with theories from the geological sciences. Ecological processes have a dramatic effect on the distribution of biologically-important chemicals (e.g., C, N, P, O). Over billions of years, ecological interactions have transformed the planet from an environment with high ultraviolet radiation, low availability of oxidizing and reducing ion pairs, and few energy sources, to an environment with low ultraviolet radiation, high availability of oxidizing and reducing ion pairs, and a diversity of energy sources (Burke and Lauenroth Chapter 11). Moreover, the interplay between the biotic portions of the environment with matter and energy at all levels of the ecological hierarchy (e.g., individuals, populations, communities) creates dynamic interacting systems (e.g., organismal systems, population systems, ecosystems) that can play out over global scales (Peters et al. Chapter 12). This framework makes clear that the study of ecosystems (i.e., community systems)

is an integral part of the domain of ecology. Ecosystems arise from the interactions of communities with matter and energy, and the resultant cycles, flows, and pools or standing stocks, are consequences, to at least some extent, of the activities of the biota.

Implicit in our definition of ecology is an emphasis on spatial and temporal perspectives. As elucidated in many of chapters in this book, ecological relationships can vary as a function of the scale of space or time, and environmental drivers that strongly affect variation at one scale may be markedly different than those at other scales. Consequently, ecological understanding, especially predictive understanding, is a challenge when the form or parameterization of a relationship may differ across scales, or when the identity of the dominant driver of a pattern changes with scale. One of the central challenges in ecology is the development of theories and models that integrate across levels in the biological hierarchy (Figure 15.1). In this book, the only theories that explicitly integrate across the hierarchy are metacommunity theory (Leibold Chapter 8), succession theory (Pickett et al. Chapter 9), and island biogeography theory (Sax and global change theory (Peters et al. Chapter 12), which integrates from individuals to the biosphere. Other chapters (e.g., Fox et al. Chapter 13; Colwell Chapter 14) hint at such integration, but do not explicitly model it.

The concept of hierarchical levels as applied to ecology (Odum 1971) has long been recognized in many realms of the natural and social sciences, with important philosophical considerations (e.g., Novikoff 1945; Feibleman 1954; Greenberg 1988). In addition to facilitating communication and classification in ecology, the integrative levels of organization in that hierarchy suggests that moving from individuals to communities involves increases in complexity, and that properties at higher levels can emerge from lower levels . At each level in the ecological hierarchy, emergent characteristics manifest that cannot be predicted or fully understood based on just the patterns and processes at lower levels. Moreover, interactions are horizontal, among entities at the same level (e.g., different species interact within a community) as well as vertical. In addition, influences are fully complementary, in that processes at higher levels can affect properties at lower levels (e.g., communities affect populations). For these reasons, both reductionist and system approaches to ecological understanding are by themselves insufficient and sometimes misleading. In ecology, an inability to clearly and unambiguously identify the spatiotemporal limits of entities at each hierarchical level (e.g., individual,

population, community) may conspire to further challenge the development of predictive understanding. For example, we may find it easier or less arbitrary to distinguish unitary individuals than to distinguish unitary communities, making it progressively more complicated to understand the linkage between pattern and process as we scale the ecological hierarchy.

THE DEVELOPMENT OF THEORY

The chapters in this book represent a diversity of ecological theories that differ greatly in content and scope, as well as in their degree of theoretical maturation. The chapters also differ in the extent to which their syntheses summarize, clarify, amplify, integrate or unify theoretical constructs to the advancement of ecological understanding. In Chapter 1 we presented a hierarchical view of theory consisting of three tiers: general theories, constitutive theories, and models (Table 1.1). We noted, however, that this division into three tiers was arbitrary. The hierarchy is really a continuum. About half of the chapters primarily focus on the model end of this continuum because they provide explicit directions for the building of models: foraging theory (Sih Chapter 4), niche theory (Chase Chapter 5), population dynamics theory (Hastings Chapter 6), enemy-victim theory (Holt Chapter 7), island biogeography theory (Sax and Gaines Chapter 10), and ecological gradient theory (Fox et al. Chapter 14). The others – metacommunity theory (Leibold Chapter 8), succession theory (Pickett et al. Chapter 9), ecosystem theory (Burke and Lauenroth Chapter 11), global change theory (Peters et al. Chapter 12), and biogeographical gradient theory (Colwell Chapter 14) – are much more like general theories because their propositions are closer in nature to fundamental principles and define the domain of their models rather than provide explicit rules for model building. So, even for the constitutive theories presented in this book, additional constitutive theories could be developed that are either narrower in scope and act to unify some particular set of models or are broader in scope and aim at uniting other constitutive theories.

To some extent the tendency of a chapter to be at one end or the other of that continuum from general to specific-model is a function of the maturity of the theory. Foraging theory, population dynamic theory, enemy-victim theory, and island biogeography theory are all quite mature and the authors of those chapters focused on models. In contrast, metacommunity theory, global change theory, and biogeographical gradient theory are less mature and those authors presented more general views of their topics. In some cases, the very maturity of the theory in combination with the complexity of the domain led to a chapter that was more general in focus (succession theory and ecosystem theory). In other cases, the theories have arisen out of attempts to synthesize across competing models, leading to a more model-focused approach of a less mature theory (niche theory and ecological gradient theory).

Theories may assume a number of different roles (R. Creath, unpublished ms.). They represent generalizations that extend the scope of the particular data that espoused those generalizations. They generate concepts that extend beyond what can be expressed in observation alone. In these two capacities, theories are primarily descriptors of the world. In addition, theories are a framework for guiding and evaluating research, paradigms *sensu* Kuhn (1962) or research traditions *sensu* Laudan (1977). To a greater or lesser extent, all of the constitutive theories in this book play each of these roles.

Many controversies arise within a domain because of a failure to differentiate between a core concept that is neutral and broad, versus various incarnations of that concept, that may be narrow and specific. The differences among particular models allow the more general theory to be broadly applicable, depending on circumstances defined by the distinguishing assumptions of those models. Understanding the features that favor one model over another thereafter becomes a unifying strength of the theory to account for myriad empirical observations. The framework for understanding disturbance and succession (Pickett et al. Chapter 9), exemplifies this process. The frameworks for niche theory (Chase Chapter 5), enemy-victim theory (Holt Chapter 7), and metacommunity theory (Leibold Chapter 8) perform similar unifications.

Roles of theory in ecology

The body of theory in ecology accounts for our observations about the natural world and gives us our predictive understanding through the use of models. It organizes those models into constitutive domains that provide a robust intellectual infrastructure. That organizational framework provides a blueprint of the strengths and weaknesses of our understanding, motivating future empirical and theoretical work, and catalyzing research agendas. Refinements of theory can identify the mechanistic bases of patterns and processes about which we have considerable confirmation, as well as distinguish the ideas and relationships that are in flux or about which there is considerable uncertainty.

In the title to his chapter, Kolasa (Chapter 2) emphasizes that "theory makes ecology

evolve." Using a historical perspective, he identifies the genesis of the ideas that formed the bases of the fundamental principles of the domain of ecology (Table 1.3). He forcefully argues that theory and empiricism are inextricably intertwined, not mutually exclusive undertakings, and that theory is a vehicle for sharing knowledge across domains as well as for targeting efforts to fruitfully deepen or broaden the scope of ecological understanding. This too is the broad perspective that can be gleaned from other chapters in this book. It is a basis for our contention that the proclamation about the death of theory (Anderson 2008) is fatally flawed, even in an age of cyberinformatics. Given the deluge of data, theory helps to separate data that are irrelevant to a particular domain of interest from that which advances understanding. Theory helps to organize multiple lines of evidence in an efficient manner. Theory provides connections among ideas and concepts within domains, and among ideas and concepts from different domains. Theory provides insights into new data requirements needed to distinguish among or resolve differences among competing views of the world.

Odenbaugh (Chapter 3) clarifies the nature of a unifying theory, and reinforces Kolasa's contention that models alone, no matter how mathematically elegant or predictive, are incomplete aspects of a mature theory. Odenbaugh challenges ecologists and evolutionary biologists to further integrate disciplinary understanding with a goal of exposing the spatiotemporal interdependence of ecological and evolutionary processes: current ecological processes are in play because of past evolutionary processes and current evolutionary processes are in play because of past ecological processes. He illustrates past efforts in this area by focusing on the work of MacArthur and his collaborators. Odenbaugh argues that although their ideas were formative and stimulatory to generations of ecologists, they did not succeed in unifying ecology, despite their intentions. Rather, their work provided the discipline with natural selection thinking, a focus on model building, and a strategy aimed at predictive understanding and generality, instead of only descriptive understanding, as in natural history. This same tension between predictive understanding and descriptive natural history helped shape the origins of ecology as a discipline at the beginning of the twentieth century (Hagen 1992). As evidenced by the chapters in our book, this striving towards predictive understanding continues to shape the discipline.

Multicausality

Ecological systems have a critical property – multicausality – that affects the structure and evaluation of ecological theories (Pickett et al. 2007). In general, multicausality (Figure 15.2A) occurs when more than one driving factor (Xs in figure) effects an outcome (Y in figure). For heuristic purposes, we distinguish a number of general types of multicausality. First are instances where variation in a particular characteristic arises as a consequence of variation in only a subset of the possibilities driving factors (Figure 15.2B). For example, each of three factors (e.g., X₁, X₂, or X₃) could affect an outcome, but they do not all do so in concert in all circumstances. In some circumstances, only X₁ and X₂ might effect the outcome whereas in other circumstances, only X₁ might do so. From an analytical perspective, this results in multivariate causation in the former situation and univariate causation in the latter situation. Second are instances where all of the driving factors in concert effect an outcome (Figure 15.2C). The multiple driving factors could act together in an additive manner or could do so in complex ways that are non-linear and include direct and indirect effects.

These aspects of multicausality are important for the structure of ecological models and their evaluation. If a model includes all of the multiple causes, it will provide robust predictions or explanations. For multicausal models of the first type (Figure 15.2B), it is necessary that the model include only the particular driving factors acting in a particular situation, but the causes that are included in the model may differ from situation to situation. Thus, it is not necessary to know all of the possible causes in all situations, just the ones that are important in the situation under consideration. For models of second type (Figure 15.2C), it is necessary to know all of the driving factors in advance of model construction.

For all types of multicausality, if a model does not include all causes, the utility of the model depends on whether those causes have additive or non-additive (i.e., interactive) effects on the outcome. When effects are additive (Figure 15.2B), conclusions about the relative magnitudes of the processes included in the model are robust. The excluded factors may affect absolute predictions of a model, but not relative ones. On the other hand, if the causes interact (Figure 15.2C), then the magnitudes and rank orders associated with one causative mechanism may depend on those of another. At minimum, it is necessary to acknowledge that this is an inherent assumption in particular models.

In this book, we find both types of multicausal models. Models that deal with independent, additive causes are exemplified by niche theory (Chase Chapter 5), population

dynamics theory (Hastings Chapter 6), enemy-victim theory (Holt Chapter 7), island biogeography theory (Sax and Gaines Chapter 10), and ecological gradient theory (Fox et al. Chapter 13). Interactive causes are notable in foraging theory (Sih Chapter 4), metacommunity theory (Leibold Chapter 8), ecosystem theory (Burke and Lauenroth Chapter 11), global change theory (Peters et al. Chapter 12), and biogeographic gradient theory (Colwell Chapter 14). Perhaps the most extreme version of such interactions is found in succession theory (Pickett et al. Chapter 9).

When evaluating models, the two types of multicausality and the details of their interactions have important implications for how an experiment would be designed. For the first type of multicausality with additive effects, rigorously holding constant all factors other than those under investigation would be most informative. Even with non-additive effects, an experiment would manipulate just a few factors. Most laboratory and greenhouse experiments are of these types. For the second type of multicausality, unless one knew all of the necessary causes and their interactions, a field experiment would be more informative. Although one or a few factors might be deliberately manipulated, other necessary factors would be free to also contribute. Importantly, statistical techniques such as structural equation modeling, which are capable of identifying causal factors and of incorporating direct and indirect effects (Grace 2006), could be employed with non-experimental data.

Because of the second type of multicausality, some philosophers of science conclude that we can never determine the true explanation of a phenomenon because multiple alternative explanations always exist (Suppe 1977). In practice, ecologists must often use multiple lines of evidence to discern the relative roles of ecological processes in producing patterns (e.g., Carpenter 1998). See Scheiner (2004) for a more complete discussion of the use of total evidence in ecology.

Another aspect of multicausality is that some causal processes are proximate and others ultimate. Consider the question: Why are male lions larger than female lions? A proximate explanation involves development and food intake during growth. A more ultimate explanation involves sexual selection: larger males are better able to monopolize a group of females. Beyond those processes may be phylogenetic effects involving all felids or carnivores. These alternative explanations often derive from different general theories, so a given constitutive theory needs to either draw on those multiple general theories or should acknowledge the limitations of its explanatory scope.

Spatial variation, temporal variation, and scale

Environmental heterogeneity, both abiotic and biotic, is core to ecological processes as shown by its prominence in the theory of ecology (Table 1.3, principles 3, 5 and 6). This heterogeneity creates a central role for the importance of scale in ecological theories.

Geographic space and ecological space are intimately intertwined. This intersection can be seen most clearly in two theories. Biogeographic gradient theory (Colwell Chapter 14) presents a synthetic framework for the creation of a theory of spatial gradients (e.g., latitude, elevation, depth) that operate at broad geographic scales. Broad-scale patterns of species richness and range size are an emergent property arising from the sum of species-specific responses. Metacommunity theory (Leibold Chapter 8) bridges local and regional scales (i.e., mesoscale ecology) by considering the extent to which local filters and dispersal determine the composition and species richness of sets of communities.

Interactions of ecological processes can change over space and time. Within a single community their relationships change as a result of disturbance and succession (Pickett et al. Chapter 9). Those interactions are now relatively understood as befits a theory that has been developing since the origins of ecology in the late nineteenth century (Cowles 1899; Clements 1916). At the other end of the spatial and temporal scale are those global changes initiated by human activities (Peters et al. Chapter 12). Borrowing concepts from hierarchy theory (Allen and Starr 1982) and landscape ecology (e.g., Peters et al. 2006; Peters et al. 2008), global change theory addresses issues associated with the consequences of large scale, human initiated disturbances such as global warming, urbanization, and agricultural intensification. This theory is implicitly scale-sensitive, suggesting that fine-scale relationships between pattern and process interact with broad-scale relationships resulting in spatial heterogeneity and differential connectivity among spatial units.

Conservation, management, and policy

Ecology as a discipline and ecologists as scientists have changed greatly from the middle of the last century when the Nature Conservancy was formed. That organization was founded by a

group of ecologists who were frustrated with their inability to get the leadership of the Ecological Society of America (ESA) to address the practical and policy implications of their science. In contrast, today the ESA has taken a leadership role in translating science into policy.

It is telling, though, that such linkage is mostly absent from this book, despite our instructions to authors that they should address those issues. Only two chapters do so explicitly: island biogeography theory (Sax and Gaines Chapter 10) and global change theory (Peters et al. Chapter 12). It is not surprising that these chapters address those concerns. Island biogeography theory has long been entwined in efforts to determine the best design for nature reserves (Burgman et al. 2005), especially the SLOSS ("Single Large Or Several Small") debate of the 1980s. Today, global change has become a central focus of both science and public policy. The magnitude and rate of change are both great, and these anthropogenically induced changes will likely affect all levels in the ecological hierarchy, often in dramatic ways and likely over broad spatial extents.

That is not to say that the other theories in this book are not also relevant to applied issues. For example, population dynamics theory (Hastings Chapter 6) is used extensively for population viability analyses. Similarly, enemy-victim theory (Holt Chapter 7) is useful in understanding pathogen-host interactions in agricultural settings as well as the dynamics of infectious diseases as they relate to public health. Rather, we ecologists tend to separate theory development from theory application. The drive for theory development often comes from basic research questions, with application and additional refinement of theory coming later. Global change theory is a notable exception. Its impetus arises from current concerns about where our planet is headed as a consequence of anthropogenic contributions to greenhouse gases and expansive modifications of landscape structure and configuration throughout the world.

Much of the application of theory to questions of management has focused on optimization issues (i.e., maximum sustained yield) related to production of particular agricultural crops or harvests of particular species of wildlife for human consumption or use. A more holistic approach that considers management from an integrated, multispecies ecosystem perspective is gaining ground because of its ability to include both direct and indirect effects on targeted species, the species with which they interact, and the ecosystem services that they provide to humans (Peterson 2005). In many ways, this heralds the emergence of a new scientific discipline – socioecology – at the intersection of the social sciences, environmental sciences, and

engineering. This new discipline explicitly considers coupled human and natural systems as the domain of interest (Figure 15.3). At its foundation is the theory of disturbance and succession, with human social systems as both the drivers of and respondents to change. These dynamic feedbacks must be used in policy decisions if they are to lead to adaptive management with a goal of enhancing resilience and long-term sustainability. Such a theory would focus on cycles of disturbance and recovery (succession) within the socioecological system (including its biotic and abiotic constituents), and would rely on an understanding of successional dynamics (Pickett et al. Chapter 9) and ecosystem function (Burke and Lauenroth Chapter 11). In so doing, it considers humans as ecological engineers or drivers of change (disturbance agents) that affect landscape configurations of local ecosystems, each with positive or negative consequences to human well being. Moreover, it considers human well-being as providing feedback to human actions via policy and management. Perhaps the greatest challenge to face society and science in the twentyfirst century will be developing a predictive understanding of coupled human and natural systems – socio-ecological systems – so that policy and management can be responsive to longterm goals of sustainability. The further development of theory across all of ecology will play a critical role in the ultimate success of such an endeavor.

INTEGRATION AND UNIFICATION: THE FUTURE OF ECOLOGY

Despite the different levels in the biological hierarchy that the chapters in this book address, as well as particular interactions that form the focus of their expositions, they share a number of common features. Each chapter defines a suite of basic propositions within a particular domain, and combines different state variables or parameters in alternative ways to provide understanding or prediction about central ecological phenomena. Each links the propositions associated with its domain back to the fundamental principles of ecology (Table 1.3). The chapters identify central models within their domains; some are conceptual while others are more precise and mathematical in nature. As such, these chapters consolidate the state of understanding and accelerate the process of theory unification. In addition, each chapter clarifies connections between its focal domain and the domains of other chapters or sub-disciplines of ecology, enhancing integration. Various chapters illustrate how different assumptions lead to different models. A failure to substantiate a particular model does not necessarily mean that the more general theory with which it is associated is wrong or useless. Rather much of ecology deals with
understanding the conditions that favor one model over another, and how these conditions relate to the formalized assumptions of each model.

During discussions at the workshop that preceded this book and during the process of articulating the various constitutive theories, a common claim was that one person's theory was central to all of ecology and that all other domains could be viewed as ancillary to her or his domain. Such viewpoints are to be expected as we attempt to build a set of integrated constitutive theories. Our general theory tells us that the constitutive theories must be linked to each other. As noted in the various chapters, each constitutive theory links directly with numerous other theories. In some cases the theories share similar propositions. Some of this sharing is expected and obvious (e.g., population dynamics theory (Hastings Chapter 6) and enemy-victim theory (Holt Chapter 7); ecological gradient theory (Fox et al. Chapter 13) and biogeographical gradient theory (Colwell Chapter 14). In other cases, overlaps become more apparent after propositions are formalized, for example the role of connectivity in metacommunity theory (Leibold Chapter 8) and global change theory (Peters et al. Chapter 12). In yet other cases, the propositions of one theory can point to ways that other theories can be modified, for example the role of species interactions in niche theory (Chase Chapter 5) as a guide to adding such interactions to island biogeography theory (Sax and Gaines Chapter 10) or ecological gradient theory (Fox et al. Chapter 13). No single constitutive theory is at the center of ecology. Rather all are connected and overlap to some degree and together define the science of ecology.

A more comprehensive unification and integration of ecology would be advanced by applying these approaches to domains of ecology beyond those considered in this book (e.g., competition, mutualism, food webs, and landscapes). For example, the metabolic theory of ecology (West et al. 1997; Brown et al. 2004) currently consists of a single model that has been applied to a variety of questions (e.g., Allen and Gillooly 2009). Progress would be furthered through the articulation of the propositions that underlie that model coupled with an attempt to develop alternative models derived from those same propositions. Such alternative models would help to clarify the debate around this theory (e.g., Hawkins et al. 2007; del Rio 2008). It would lead to the testing of alternative hypotheses, going beyond the current practice of simply fitting data to a single model, as is frequently done across all of ecology.

We do not mean to imply that ecology will advance or mature only via a single approach, such as that advocated in this book. Indeed, understanding within a domain evolves via a variety of activities. Such a diversity of approaches can lead to robust formulations of the intellectual framework – the theory – that distinguishes ecology, integrates its components, and identifies lacunae in understanding or shortcomings in empirical validation.

The past 50 years in ecology have seen the development of two contrasting approaches to model development. One is the ecosystem approach, exemplified by the work of Odum and his collaborators (Odum 1971), that attempts to build models that are highly complex and specific. The other is the evolutionary ecology approach, exemplified by the work of MacArthur and his colleagues (Odenbaugh Chapter 3), that aims to build very simple and general models. Although often seen as antithetical (Odenbaugh 2003), the theoretical framework presented in this book can encompass both modeling approaches (e.g., Holt Chapter 7 and Burke and Lauenroth Chapter 11). The challenge for all modeling approaches is to determine the underlying propositions that provide the theoretical framework for a set of models. For simple, general models the move to more general propositions is relatively straightforward, although still not a trivial exercise as was discovered by the contributors to this book. For complex, specific models deriving general propositions is less straightforward. Recent advances in structural equation modeling (Grace et al. (in press)) provides one avenue by which such models can be united within a general framework.

In summary, the process of unification and integration is well underway within the various domains of the constituent theories of ecology, as well as at the level of the principals of ecology in general, including its integration with the rest of the biological disciplines (Scheiner submitted). An uber-theory, in the sense of an all encompassing model or mathematical formulation, is unlikely to characterize ecology in its full diversity of content based on the hierarchy of interacting systems.

We are hopeful that from these modest beginnings, advances in ecological understanding will be accelerated by a faithful and consistent application of integrative and unifying approaches to the development of theory, such as those considered in this book. We trust that these chapters will strengthen the foundations of ecological understanding, and help to herald a time of an intensified interest in ecological theory. We are not viewing the death of theory. Borrowing from Winston Churchill (10 November 1942), "... this is not the end. It is not even the beginning of

the end. But it is, perhaps, the end of the beginning" of a revitalization in the advancement of theory as a vehicle for promoting deep understanding of ecological systems.

Acknowledgments

The book by Pickett, Kolasa, and Jones (1994), *Ecological Understanding: The Nature of Theory and Theory of Nature*, was seminal to the completion of this chapter. In addition, many of the ideas we develop arose from interactions among the chapter authors who were participants in a workshop on the Theory of Ecology, supported by the University of Connecticut through the Center for Environmental Sciences and Engineering and the Office of the Vice Provost for Research and Graduate Education. Support to MRW was provided in part by the National Science Foundation via grant DEB-0614468. This work was done while SMS was serving at the U.S. National Science Foundation and on sabbatical at the Center for Environmental Sciences and Engineering at the University of Connecticut. The views expressed in this chapter do not necessarily reflect those of the National Science Foundation or the United States Government.

Literature cited

- Allen, A. P., and J. F. Gillooly. 2009. Towards an integration of ecological stoichiometry and the metabolic theory of ecology to better understand nutrient cycling. Ecology Letters 12:369-384.
- Allen, T. F. H., and T. B. Starr. 1982, Hierarchy: perspectives for ecological complexity. Chicago, IL, University of Chicago Press.
- Anderson, C. 2008. The end of theory: the data deluge makes the scientific method obsolete, Wired Magazine.
- Brown, J. H., J. E. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771-1789.
- Burgman, M. A., D. B. Lindenmayer, and J. Elith. 2005. Managing landscapes for conservation under uncertainty. Ecology 86:2007-2017.
- Carpenter, S. R. 1998. The need for large-scale experiments to assess and predict the response of ecosystems to perturbation, Pages 287-312 in M. L. Pace, and P. M. Groffman, eds. Successes, Limitations, and Frontiers in Ecosystem Science. New York, NY, Springer.
- Clements, F. E. 1916, Plant Succession. Washington, DC, USA, Carnegie Institute of Washington.
- Cowles, H. C. 1899. The ecological relations of the vegetation on the sand dunes of Lake Michigan. Botanical Gazette 27:95-117, 167-202, 281-308, 361-391.
- del Rio, C. M. 2008. Metabolic theory or metabolic models? Trends in Ecology & Evolution 23:256-260.
- Feibleman, J. K. 1954. Theory of integrative levels. British Journal for the Philosophy of Science 5:59-66.
- Grace, J. B. 2006, Structural Equation Modeling and Natural Systems. Cambridge, UK, Cambridge University Press.
- Grace, J. B., T. M. Anderson, H. Olff, and S. M. Scheiner. (in press). Improving the connection between biological data and theory through the use of structural equation meta-models. Ecological Monographs.
- Greenberg, G. 1988. Levels of social behavior *in* G. Greenberg, and E. Tobach, eds. Evolution of Social Behavior and Integrative Levels. Hillsdale, NJ, Erlbaum.

- Hagen, J. B. 1992, An Entangled Bank : The Origins of Ecosystem Ecology. New Brunswick, NJ, Rutgers University Press.
- Hawkins, B. A., F. S. Albuquerque, M. B. Araújo, J. Beck, L. M. Bini, F. J. Cabrero-Sanudo, I. Castro-Parga et al. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877-1888.
- Kolasa, J., and S. T. A. Pickett. 1989. Ecological systems and the concept of biological organization. Proceedings of the National Academy of Sciences 86:8837-8841.
- Kuhn, T. S. 1962, The Structure of Scientific Revolutions. Chicago, University of Chicago Press.
- Laudan, L. 1977, Progress and Its Problems: Toward a Theory of Scientific Growth. Berkeley, University of California Press.
- Novikoff, A. B. 1945. The concept of integrative levels and biology. Science 101:209-215.
- Odenbaugh, J. 2003. Complex systems, trade-offs and mathematical modeling: a response to Sober and Orzack. Philosophy of Science 70:1496–1507.
- Odum, E. P. 1971, Fundamentals of Ecology. Philadelphia, PA, W. B. Saunders.
- Peters, D. P. C., B. T. Bestelmeyer, J. E. Herrick, H. C. Monger, E. Fredrickson, and K. M. Havstad. 2006. Disentangling complex landscapes: New insights to forecasting arid and semiarid system dynamics. BioScience 56:491-501.
- Peters, D. P. C., P. M. Groffman, K. J. Nadelhoffer, N. B. Grimm, S. L. Collins, W. K. Michener, and M. A. Houston. 2008. Living in a connected world: A framework for continental-scale environmental science. Frontiers in Ecology and the Environment 5:229-237.
- Peterson, G. D. 2005. Ecological management: control, uncertainty, and understanding *in* K. Cuddington, and B. Beisner, eds. Ecological Paradigms Lost: Routes of Theory Change. Burlington, MA, Elsevier Academic Press.
- Pickett, S. T. A., J. Kolasa, and C. G. Jones. 2007, Ecological Understanding: The Nature of Theory and the Theory of Nature. New York, NY, Elsevier.
- Scheiner, S. M. 2004. Experiments, observations, and other kinds of evidence, Pages 51-71 *in* M.
 L. Taper, and S. R. Lele, eds. The Nature of Scientific Evidence. Chicago, IL, University of Chicago Press.
- —. submitted. Towards a conceptual framework for biology. Quarterly Review of Biology.

- Suppe, F. 1977, The Structure of Scientific Theories. Urbana, Illinois, USA, University of Illinois Press.
- West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. Science 276:122-126.

Figure 15.1. The biological organizational hierarchy extends from molecules to biomes. At each level, biological entities (e.g., cells, individuals, communities) interact with matter and energy (double headed horizontal arrows) to form living systems. The domain of ecology is defined by interactions at the level of individuals and higher (signified by black rather than gray lettering), and is characterized by an organismal perspective (signified by the shaded ellipse at the individual level). Each level in the biological hierarchy is associated with other levels (e.g., populations comprise individuals of the same species and populations of different species comprise communities) in the hierarchy (represented by vertical arrows on the left). Additional connectivity among levels occurs because the matter and energy that fuel the activities of all biological entities flow and cycle, respectively, in biological systems regardless of level in the hierarchy (indicted by vertical arrows on the right). Moreover, then nature of the ecological entities and their interactions changes over time as a consequence of evolution, resulting in complex dynamics and multiple feedbacks. (modified from Odum 1971).



Figure 15.2. These diagrams represent various types of multicausality, a situation in which more than one driving factor (represented by Xs) effects variation in focal characteristic (represented by Y). (A) A general model that defines the candidate driving factors (solid grey lines) and the responding characteristic of interest. (B) Only a subset of the possible driving factors have an effect depending on particular circumstances (represented by solid arrows). In this case, the model is additive and multivariate, but under other circumstances it could be univariate. (C) All of the candidate driving factors are interactive. Even more complex situations can arise, where only subsets of the candidate driving factors come into effect in particular circumstances, and the factors interact in a non-additive way, including direct and indirect effects.



Figure 15.3. This conceptual model illustrates key linkages between natural and human systems that together constitute a socioecological system. It recognizes (1) that the functionality of natural systems varies along a continuum from intact to degraded, with each providing inputs (both positive or negative) to human well being; (2) that human decisions affect ecosystems positively, via recovery, restoration or reclamation, or negatively via degradation; and (3) that real or perceived well being of humans should directly feedback on decision making (policy) so as to adaptively and sustainably manage natural systems.

