Introduction
Chap 1. The Principles of Ecological Integration

TOWARD
A UNIFIED ECOLOGY

TIMOTHY F. H. ALLEN
AND THOMAS W. HOEKSTRA

WITH ILLUSTRATIONS
BY JOYCE V. VANDEWATER

SECOND EDITION

COLUMBIA UNIVERSITY PRESS   NEW YORK

Columbia University Press
Publishers Since 1893
New York   Chichester, West Sussex

cup.columbia.edu

Library of Congress Cataloging-in-Publication Data

Allen, T. F. H.
Toward a unified ecology / Timothy F.H. Allen and Thomas W. Hoekstra. — Second edition.
pages cm. — (Complexity in ecological systems series)
Includes bibliographical references and index.
ISBN 978-0-231-16888-5 (cloth : alk. paper)—
ISBN 978-0-231-16889-2 (pbk. : alk. paper)—ISBN 978-0-231-53846-6 (ebook)
1. Ecology—Philosophy. I. Hoekstra, T. W. II. Title.
III. Series: Complexity in ecological systems series.

QH540.5.A55 2015
577.01—dc23
2014033202

Columbia University Press books are printed on permanent
and durable acid-free paper.
This book is printed on paper with recycled content.
Printed in the United States of America
c 10 9 8 7 6 5 4 3 2 1
p 10 9 8 7 6 5 4 3 2 1

COVER DESIGN: Noah Arlow

CONTENTS

3. THE ECOSYSTEM CRITERION 112

4. THE COMMUNITY CRITERION 156

5. THE ORGANISM CRITERION 208

PREFACE

This second edition of *Toward a Unified Ecology* comes some twenty years after the first. Since the first edition, notions of complexity theory have become current. The first edition was ahead of its time, but this new edition is able to take advantage of more recent sources that were not available in the early 1990s. We have since been involved with two books that pressed ideas of complexity and hierarchy theory. Ahl and Allen presented a slim book in 1996, *Hierarchy Theory: A Vision, Vocabulary and Epistemology*. That was the condensed version of Allen and Starr's 1982 *Hierarchy: Perspectives for Ecological Complexity*. In 2003, with Joseph Tainter, we published *Supply-Side Sustainability*, one of the few books to integrate social and ecological science into a workable whole. We have taken advantage of that integration in this new edition. There was a set of principles there, and we have folded them into the management sections of the new edition.

Another important idea in our work in the last decade has been notions of high gain and low gain. Ecologists think about using resources, but focus in particular on running out of resource. By contrast, economists know we do not run out of anything; things just get more expensive. As a result, economists watch their systems adapt, while ecologists cry doom, gloom, and extinction. The idea of profit in using resource is alien to ecology, but is well embedded in notions of high and low gain. Gain is profit. High gain and low gain relate to r versus K selection in ecology, but there is more generality in gain. For instance, high gain predicts system behavior from flux driving the system. Low gain looks at constraints as predictors. The high/low distinction pits rate dependence against rate independence, and notes that the two classes of explanation cannot be readily applied at once. If you can see the river flowing, you cannot see the effect of the dam; if you can see the dam creating a lake, you cannot see the river (a wider purview would see them both, but the rate dependence of the river would be lost in the ribbon). To an extent, high versus low gain is a way of cleaving apart levels of analysis. Gain is the complexity science version of r and K selection, where K gives structure and r is the dynamics that links structure from different levels.

We were committed to narratives in the first edition, but could not be explicit then for fear of out-and-out rejection. Narrative is now more acceptable in science. Storytelling was always there, but was not recognized; it was used as a cheat in the realist world of modernist science. The power of narrative is now accepted, but the persistent overt realism in the mainstream of ecology has been unhappy that narratives are not about truth. Stories are statements of points of view. Focus on verity and truth is associated with entrenched modernism, which insists that science approaches truth. It may or may not, we can never know. There are subtleties here, such as true and false versus not true and not false.

Reflexive realism will be frustrated at our position because we are denying some of its premises. But we ask for tolerance; the reason is that we only want to rein in realist excesses, not deny realism. In the end, we are realists too, but ours is a realism that is so measured as not to be an easy target in more sophisticated logical discourse. Our realism does not invite methods of investigation that get muddled. We say to realists, we are on your side. But we do want to modify your claims about reality so they are as fully valid as they can be in the final analysis. The device of science from its inception is to investigate what appears before us, and indeed what appears to be true. But scientific investigation always aims at showing our first impressions to be untrue. Investigation never proves anything; it just shows that such and such cannot be true. Science moves by disproving things, not by proving anything. And science shows appearance at first glance to be wrong and illogical time and time again, such that it would be unwise to expect anything else. Asserting truth would seem reckless, at least until the investigative process is seen as concluded.

To press the point home, we note that science is even detecting at a metalevel that important reliable findings are incorrect across a range of fields. Jonathan Schooler has reported the decline effect.[1] That work shows significance and reliability to decline linearly as work is repeated by other scientists. Perhaps feeling that parapsychology had it coming, many will be delighted that the effect was first noted in the 1930s, where statistically reliable results in parapsychology were shown to decline in attempts to validate early work. But not so fast; the decline effect has been reported in a number of fields, tellingly in the effectiveness of drugs for mental health. Drugs such as aripiprazole (e.g., Abilify) were at first shown to be effective with robust statistical significance, but attempts to reproduce those effects decline some 30 percent in significance each time. Decline effect has been shown in ecology with regard to symmetry of tail feathers and mating success in birds.[2] Reliability of results declines faster in popular fields, where repetition is most likely to occur.[3] If well-performed experiments, accepted in peer review, achieving the highest professional standards cannot be validated and reproduced, it would seem that a certain caution should be applied to truth and reality with regard to science.

Science is designed to be a naysayer. So assertions at the outset that such and such is true enough to be used as a benchmark are unwise. Such insistence inserts a point of reference that exactly blunts science as a device. Science is as good a device as we have, so keeping it honed is important. The danger with hasty realism in science is that it interferes with doing scientific measurement and testing. In the end, the authors here are hard bitten, driven by data and experience; there is nothing airy-fairy about our posture, despite our proclivity for abstraction. We were both empiricists in our training and early careers. All we are saying is that using truth as a reference at the base of measurement and provisional notions is unwise. We can never know ultimate truth, so it is not a reliable reference; science itself shows us that fact time and again. As scientists take positions, we insist that they take proper responsibility for those decisions; otherwise, they will fool themselves. But we must hasten to add that once the decision making and testing is over, and we agree the result is good, then saying, “We are now closer to reality” does no harm and has intuitive appeal. So we are not bottom-line antirealist. Reality as a conceptual device has its uses. The feeling of reality in stories is one of the appeals of narrative, even if stories are neither true nor false. We do not need to be tethered to reality as a device, but that does not stop it from being a powerful, workable context.

The organization in this new edition follows that of the first edition. Indeed, our philosophy has not changed. The introduction and chapter 1 lay out the general premise that cleaves scale from type in investigating complexity. Clarity on scale and type allows organized movement around a labyrinth that is complicated by being tiered. The tools we introduce are ways of dealing with complexity in remarkably practical ways.

The middle section of this edition works its way through types of ecology, investigating scaling issues within each type. We use the same types in the same order as in the first edition. It worked then, and we are happy to use the same framework again. Reference to new papers and recent research amplifies what we said two decades ago. So there is old and new here.

New to this edition is an explicit chapter on the use of narratives and models in ecology in particular, but also in science in general. That precision in philosophy of science informs the last two chapters, which, as in the first edition, are about management and basic science ecology. We feel the whole is better integrated in this edition.

Allen did use the first edition to teach his general ecology course. He made it work well. The book has the advantage of covering the whole discipline with an intellectual challenge. We invite others to do the same. But this is not a normal textbook. Maybe it is time to rethink how textbooks might work. There is more thinking here and fewer facts. It is sad that students like facts because facts appear concrete for memorization for the test. But one does not have to use this volume as a textbook for teaching ecology; a large part of our audience is likely to be seminars in ecology. As teachers ourselves, we are aware of something like fifteen working weeks in a semester. Here we have eleven units of text, so the number is not quite right for one chapter a week. There are several ways to extend the readings to get the fit with the weeks in a semester course. First, chapter 1 introduces material that is likely to be very new and distinctive. It might make sense to give chapter 1 two weeks of discussion: the first week for the conceptual setup, and the second week to do justice to our solution, the cone diagram. We are also aware that some of the chapters on types of ecology are large. Landscapes are given a long treatment. Ecosystems and communities are behemoths. Each could be treated first as theory and second with examples. In fact, even used as just one chapter for a whole discussion, it makes sense to emphasize the theory/example tension in the big chapters.

But everyone in ecology needs this work. It anchors us in the world when modernist mechanism is breaking down. Our book is not so much a security blanket while the old intellectual framework fidgets itself apart as it is something to hold on to as ecology melts into economics and the humanities come to stay as collaborators.

Timothy F. H. Allen
Thomas W. Hoekstra

ACKNOWLEDGMENTS

The critical role of Clyde Fasick in supporting our work was acknowledged in the first edition dedication and described in the acknowledgments of that edition, all of which are still well deserved and appropriate for the publication of the second edition. This edition relies heavily on the first, and so acknowledgments there still apply.

There were many groups who played critical roles in developing Allen's ideas in the new edition. There are so many people that they are ordered alphabetically in their respective groups. All are significant. At the outset of the focused effort to create the second edition, Allen's graduate students met weekly for a semester to discuss what should be done: Peter Allen, Marc Brakken, Julie Collins, Keith Doyon, Nissa Enos, Brian Garcia, Cassandra Garcia, Megan Pease, Steve Thomforde, and Devin Wixon. Peter Allen there resigned to the fact, saying, "Tim will write what Tim will write." That does tend to be Allen's method, but the influence of the graduate students is not to be underestimated. Another group that has been crucial in Allen's development met every Tuesday during the semester, year after year until 2012. It was called "Sandbox," and was a testing ground for many ideas. David Hall was the longest inhabitant, and attended and contributed handsomely for almost two decades. Preston Austin and Jerry Federspiel came from off campus to make wonderful Sandbox contributions. Undergraduates figured large in Sandbox, notably but not exclusively: Mike Chang, Ed Engler, David Evans, Amelia Krug, Noel Lawrence, and Anna Marie Vascan. Some Sandbox undergraduates have been wonderful coauthors with Allen on some big ideas on high and low gain and narrative: Elizabeth Blenner, John Flynn, Michael Flynn, Amy Malek, Kristina Nielsen, and Rachel Steller. Kirsten Kesseboehmer and Amanda Zellmer, as undergraduates, made huge contributions to the notions of modeling and narrative and, as authors, on some of the most important papers. Another crucial group met after Sandbox, Allen's "My engineers": Gregori Kanatzidis, Nathan Miller, Samantha Paulsen, and Edmond Ramly. They all played a role in developing the ideas for chapter 8 here. Julie Collins, Kirsten Kesseboehmer, and Noel Lawrence have backgrounds in the humanities, which they presented with unusual confidence to the scientists. Bruce Milne and his graduate student Mike Chang were generous with their new ideas on foodsheds.

Various senior scientists have been helpful as colleagues for Allen to discuss aspects of the work between editions: Thomas Brandner, Martin Burd, Steve Carpenter, Charles Curtin, Billy Dawson, Roydon Fraser, Robert Gardner, Mario Giampietro, Philip Grime, Alan Johnson, James Kay, Ronald McCormick, John Norman, Robert O'Neill, David Roberts, Edward Rykiel, John Sharpless, Duncan Shaw, Hank Shugart, Joseph Tainter, and David Waltner-Toews. Henry Horn and Tony Ives provided the anecdotes of their childhood memories. Judith Rosen was most helpful in giving the inside story on her father, Robert Rosen, and for chasing down several of his crucial quotations.

Hoekstra was fortunate to have the support of the U.S. Forest Service (USFS) in integrating his work assignments there with work on and application of *Toward a Unified Ecology*. That support allowed Hoekstra and his colleagues to test the application of concepts in the first edition. The applied research studies were carried out within the mandate for resource management by the agency and with the narrative modeling scheme described in chapter 9 of the first edition, "A Unified Approach to Basic Research." They were able to develop and apply the concepts in *Toward a Unified Ecology* from a national forest to the international scale. The results of these applied efforts support and confirm our confidence in the value of these concepts, which are expanded in this edition.

The geographic and ecological scope of these applied research studies required the involvement of many people. Unfortunately, it is not possible to individually recognize everyone involved, but only to mention the key participants. In addition, it should be mentioned that many of these studies on *Toward a Unified Ecology* concepts were also created within the context of *Supply-Side Sustainability* concepts found in our companion book with our colleague Joseph Tainter.

The most significant of *Toward a Unified Ecology* tests of concept are briefly described here. First, a 1994 workshop in Israel and its publication involved several hundred scientists and managers from more than thirty nations focused on the topic of Hoekstra and Moshe Shachak's book *Arid Land Management: Toward Ecological Sustainability*. Steve Archer, Linda Joyce, Joe Landsberg, David Saltz, No'am Seligman, and Moshe Shachak led the technical sessions of the workshop. Second, a North American applied research study in 1999 used the concepts in

Toward a Unified Ecology for developing ecological protocols to monitor forest management unit scale sustainability, the Local Unit Criteria and Indicators Development (LUCID) test. Gregory Alward, Brent Tegler, Matt Turner, and Pamela Wright worked with Hoekstra on the test that involved more than one hundred individuals in six USFS national forests and universities. In addition, there was one site in both Canada and Mexico. An applied research study using a hierarchy of ecological, administrative, and political units in 2005–2006 integrated the framework of *Toward a Unified Ecology* with social and administrative requirements. The objective was to develop local scale inventory and monitoring criteria and indicators necessary to work within the regional and national information and reporting requirements of the Forest and Rangeland Renewable Resources Planning Act. Joseph Tainter and Cedric Tyler, working with Hoekstra, led various phases of the study. There were several other applications implemented on a smaller scale.

We want to especially acknowledge the following individuals for making a special effort to provide photographs and illustrations that are used and duly credited in the book: Martin Burd, Mike Clayton, Kelly Elder, Robert Gardner, Mario Giampietro, Philip Grime, William Hereford, Alan Johnson, Ronald McCormick, Bruce Milne, Itshack Moshe, John Norman, Richard T. Reynolds, Charles Rhoades, and Monica Turner. We also deeply appreciate Matt Turner's editorial suggestions on national forest policies and protocols.

Hoekstra is doubly indebted to his wife, Joyce VanDeWater—first, for having her personal life seriously impacted by his unpredictable and challenging priorities associated with work on this book; and second, as the illustrator and partner, patiently developing and revising high-quality illustrations with us while we produced multiple versions of the book and equally as many versions of the illustrations. It is a small token of our appreciation to have Joyce identified on the title page as our partner in what is published here.

Allen remains eternally grateful to his wife, Valerie Ahl, for her huge intellectual contributions, pulled together in their 1996 book and with ongoing influence right up to the second edition. He greatly appreciates her patience and support as the work dragged on and cluttered their lives since his retirement in 2010.

TOWARD
A UNIFIED ECOLOGY

INTRODUCTION

It might seem an odd place to start, but the Beaufort scale is a combination of science and poetry. This is apt for a book on our subject because ecology is also about capturing nature and overcoming the difficulty of unambiguously communicating an overwhelming amount of information. Scott Huler wrote a splendid book, *Defining the Wind*, reporting the life of Rear Admiral Beaufort.[1] In 1806, Sir Francis Beaufort (1774–1857) was a captain in the British Admiralty. For his own use, he wrote down a set of systematized observations for wind speed.[2] One captain might write in the log that the wind was a moderate breeze, while another might call it a fresh breeze. Beaufort's words were accounts of the sails that were up on the ship: At force 0, the ship would be in full sail, but the sailors would be unable to steer because there was no flow of water past the rudder. To steer a big ship would take a force 2 wind. At force 6, a strong breeze, half the sails would be taken down in an orderly manner, with some sails only taken in halfway as opposed to stowed. Any captain would be aware of the exact conditions. At force 6, it would be "the wind to which a well-conditioned man-of-war could just carry in chase, full and by single-reefed topsails and top-gale sail." Force 7 is a moderate gale, "to which a well-conditioned man-of-war could just carry in chase, full and by double reefed topsails, jib, &c." At force 12, all sails would be down. There have been studies on past weather conditions from times long ago, based on ships logs. With the coming of steam power, sails were absent and so the description became the condition of the sea, for instance, foam streaking on the waves.

Beaufort did not pen the words that are so poetic for the wind on land; that was done in a most unlikely way. In 1906, a committee of engineers, of all things, created a lovely set of descriptions! A recitation of the Beaufort scale sounds like descriptions of a collection of Turner paintings, from idyllic to dramatic (figures 1*A* and 1*B*). Huler (2004) notes that at force 5, “Small trees in leaf begin to sway.” It is in iambic form. The second half goes on in trochaic pentameter, “Wavelets form on inland waters.” Table 1 provides a complete listing in the poetic wording of the 1906 British engineers; with two parts to most entries, it rings of a set of Japanese haiku.[3]

Beaufort and the 1906 engineers used their senses to capture a richly textured fabric of experience and standardize it so that it could be reliably communicated. Scott Huler appeared on a National Public Radio (NPR) interview, where he said he bought a small anemometer so he could calibrate his experience of the wind.[4] On a visit to a convenience store, there were some interesting wind conditions and he was frustrated when he realized that he did not have his machine. But then Huler made exactly the point we are making. He remembered the reason for his whole study: to observe conditions so that he could calibrate them by making comparisons. So he made those observations, one of which was that it was difficult to open the car door. Not as poetic as the engineers of 1906, but fully functional.

FIGURE 1*A*. Turner Painting, “Calm, from the Liber Studiorum,” 1812. Joseph Mallord William Turner, 1775–1851, Collection Tate, 178 × 267 mm.

FIGURE 1*B*. Turner Painting, “The Shipwreck,” 1802. Joseph Mallord William Turner, 1775–1851, Bridgeman Art Library, Fitzwilliam Museum, University of Cambridge, United Kingdom, 1,705 × 2,416 mm.

TABLE 1 The Beaufort Scale

0. Smoke rises vertically.
1. Direction of wind shown by smoke, but not by wind vanes.
2. Wind felt on face; leaves rustle; ordinary vane moved by wind.
3. Leaves and small twigs in constant motion; wind extends light flag.
4. Raises dust and loose paper; small branches are moved.
5. Small trees in leaf begin to sway; wavelets form on inland waters.
6. Large branches in motion; whistling heard in telegraph wires; umbrellas used with difficulty.
7. Whole trees in motion; inconvenience felt when walking against the wind; umbrellas discarded in exposed places.
8. Breaks twigs off trees; generally impedes progress.
9. Slight structural damage occurs (chimney pots and slates removed).
10. Seldom experienced inland; trees uprooted; considerable structural damage occurs.
11. Very rarely experienced; accompanied by widespread damage.
12. Countryside is devastated.

Huler's interview was a call-in show, and one of the callers reported that he had standardized his experience of the wind at Candlestick Park, the baseball stadium by San Francisco Bay, a windy place. The caller's observation that clinches his level 10 is a wind so strong it blows the foam off his beer as he watches the game—a spirited scale indeed! There has been an official update to Beaufort beyond the 1906 poetry. As of the late twentieth century, bureaucrats without much soul now reference plastic trash cans blown over, and other godless things.

The scale is used in places scattered all over the world. Canada uses it, but not the United States. Allen lived in a trailer during graduate school in North Wales. It was a rickety old thing, made of painted hardboard, not metal. He had it up on a hill with nothing substantial between him and Brazil. It was next to a hedge on the leeward side, and on the windward side it had a milk churn full of rocks with a rope over the roof that was tied to the leeward axle. The British Broadcasting Corporation (BBC) radio closed its broadcast at night with a report of sea conditions around the coast. Allen fell asleep at night listening. The places seemed romantic: "Faroe Islands: westerly, gale force 8. Scilly Isles: southwesterly, strong breeze, force 7." Allen knew he was in for a rough night if he heard "Irish Sea: westerly, storm force 10." That was as bad as it got, and the milk churn did its job.

Ecology is one of a handful of disciplines whose material study is part of everyday encounters: birds, bees, trees, and rivers. But it is a mistake to imagine that this familiarity makes ecology an easy pursuit. In fact, that is the main issue in this introduction: ease of observation makes the study harder. Being outdoors, simply enjoying nature is one thing, but a formal study is a very different matter. From the outset, we state that the very familiarity of ecological objects presents difficulties, some of the same difficulties Beaufort addressed. The familiar slips through human portals with such ease that memory is quickly overwhelmed by the sheer quantity.

To deal with the excess of human experience, ecologists must have a plan, a method of choosing significance. In any scientific study, there are the obvious steps that compress the rich experience down to a set of formal models.[5] Formal models can be equations or graphs, and may be word models. Scientists are familiar with that sort of compression. Less obvious is a prior compression down to the general area of discourse in which the model sits. This compression decides what the science is going to be about. The prior compression is less obvious than erecting the model because it is quickly forgotten as context to the second compression to the model. The first compression is taken for granted, so while it is still there, it disappears.

The first compression is to a paradigm. Thomas Kuhn identified a paradigm as a shared vocabulary, a shared methodology, and a shared view of what matters.[6] The last thing a scientific fish would discover is water, for much the same reason that devotees of a paradigm are not conscious that they have a paradigm, let alone know what theirs is. The interaction between models and paradigms is a continuing theme throughout this book. Paradigms name the things to be studied. Scientists then go on to build the narrative into which the named things fit. Paradigms are narratives; one has the story framed out so the next compression down to the model is directed at part of the narrative paradigm, filling in the gaps. Models are used to improve the quality of the narrative and press the paradigm forward. A less positive description is that a paradigm is a tacit agreement not to ask certain questions. Defense of old paradigms can be a nasty business, full of jealousy and deceit. Ecology has its fair share of paradigm fights, and we do not shrink from including the politics of them in this volume.

In natural history, there is less attention to compression, so natural historical accounts are often an accumulation of natural historical experience. A natural historical account is very narrative, but the process of improving the tale with the second compression down to models is largely neglected. In ecology, the scientist is supposed to do more than that. Failure in that responsibility since our last edition might call a lot of what passes for ecology as "natural history with numbers." We insist on theory over quantification.

Experiencing something is an act separate from deciding what that something might be called and how it might be measured. Beaufort gave nature names like "fresh breeze." The name comes from the group of equivalent wind conditions to which the experience is assigned. One hopes that when you have seen one fresh breeze, you have seen them all. The name comes from the class to which the entity is said to belong. Thus, a tree could be an organism, a plant, or a member of a particular species like maple. By being conscious of the steps toward model building, we can keep tabs on what we are doing. Beaufort makes a nice parallel exposition.

Measurement comes quite late in compressing experience into things to be investigated. You cannot measure something until a boundary for it has been decided. Before measurement, you have to have a thing to measure. That thing for Beaufort is the complex of observables about the wind. Note here that we emphasize it is *decided*; that way, we can take responsibility for our decisions. There is no excuse for abdicating our decisions to nature, as if nature could make decisions for ecological observers. No! Definitions and names do not come from nature; they come from us, just like the Beaufort scale came from Beaufort. Things may exist in the world, but they do not exist in the world as things. We create "fresh breezes," although nature makes the wind. The thing behind the names comes from human decisions that carve out a piece of the continuous experience stream so as to freeze some of nature into a thing. All sciences do this, but in ecology, the familiarity with the nature that we see might lull us into thinking that we are looking at the true nature of nature. No, we only have experience, which only later is given significance and formalized into observation.

Ecology, as with all science, is a matter of tunneling forward with decisions. Ecologists always could have tunneled their way to somewhere else, so investigators will constantly need a good record of where they have been. Formal accounts force ecologists to remember which are decisions, as opposed to which are aspects of nature. Ecologists must be aware that over time, as they find out more, the situation may change, as when Beaufort's excellent descriptions of sails became moot when sails disappeared with steamships.

Constructivist ideas about learning assert that there is no blank slate on which the world writes to human experience.[7] Remembering prior experience opens us up to recognizing certain things. Human experience is a product of prior experience. The past experience makes the present. We all see what we expect to see. With Beaufort's experience and ideas impressed on them, the 1906 engineers assisting the British Meteorological Office could easily erect new criteria for inland conditions. Beaufort opened their eyes to a fresh breeze force 5: "Small trees in leaf begin to sway; wavelets form on inland waters." Notice in constructivism, it is not the physical fresh breeze that is constructed; it is the construction of the observers' understanding and capacity for recognition (figure 2).

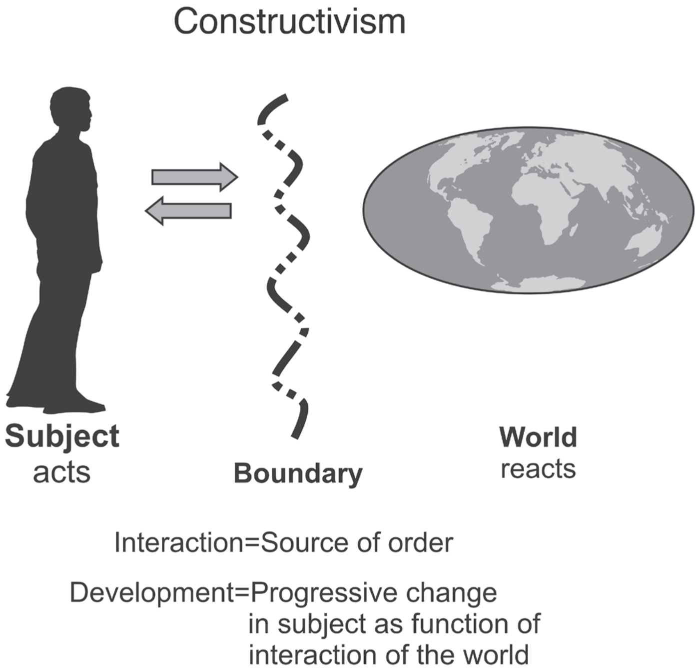

FIGURE 2. Subjects cannot experience the world itself, only sensory inputs through the filter of their senses, or through some contrivance like a pH meter. To deal with the world, the observer acts, and in some mysterious way gets an apparent reaction from the presumed world. Change occurs (sometimes imagined as progress) as the observer's understanding is constructed by interaction.

Deciding on what an ecological thing is assigns it to an equivalence class, a set of things that have something critical in common, like wavelets forming on inland waters. Classes are the organizing principle that helps human observers deal with the tidal wave of different things flooding their senses. In a step further, classes can themselves be organized. We do this intuitively much of the time, but intuition may not be an adequate organizer. At the outset, the class is simply a set of things that have something in common. But the classes that are recognized are often related one to another, whereupon the classes may become levels.

The body of theory here is hierarchy theory, and it has been particularly useful in ecology. As Eric Knox told taxonomists, since they use levels all the time (species, genera, families, etc.), it might be smart to have a theory of levels and hierarchies (figure 3).[8] Level of organization is only one sort of level, which is why we need a body of theory to keep things straight. Other sorts of levels may be based on size, that is, scalar levels. But other levels may be control levels, such as the governor on a steam engine belonging to a higher level than the machine it controls. Scale-based levels reflect the size of the things assigned to the level. Notice that in contrast, levels defined by control are not size ordered; the Watt governor is smaller than the steam engine it controls. Whatever the steam engine does, the governor

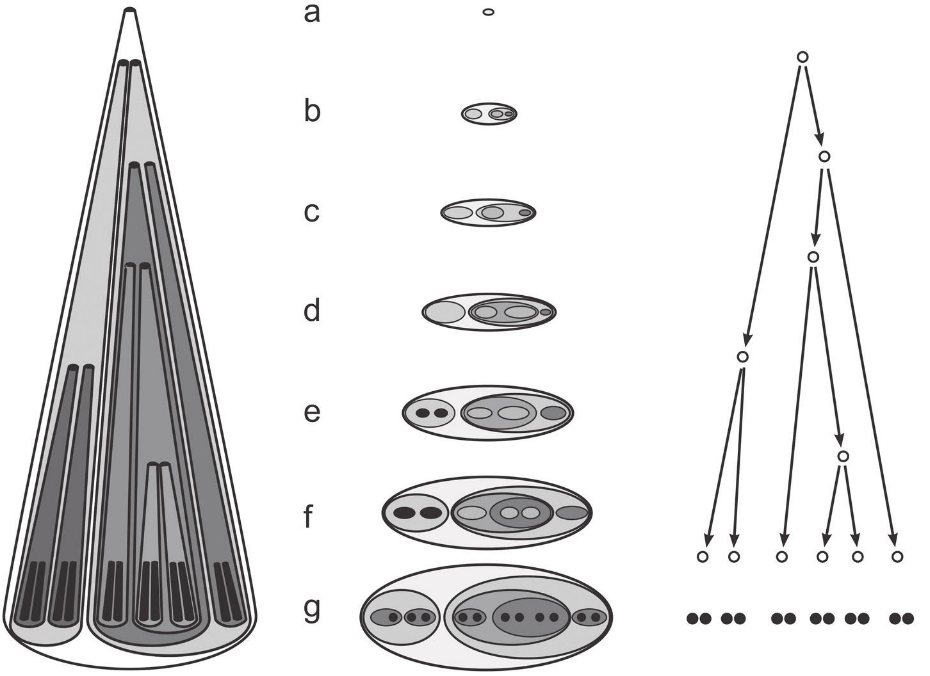

FIGURE 3. Knox (1998) shows different sorts of hierarchies that speak for themselves, but only after Knox pointed them out. The systematists he addressed were mired in what are the real hierarchies, not understanding that each hierarchy is a point of view.

has an answer; it fills the control space above the engine. Classes of larger, more inclusive things are seen as being at a higher scalar level. Scalar levels invite, but do not insist on, nesting of levels.

We have scaling tools that let us observe at different scales. To see big things at a distance, a telescope might be useful. To see small things, one might employ a microscope. Small things that move fast, like birds, might require binoculars to establish a context for the observation. Embedded in the use of a given tool are two scaling considerations: grain and extent or scope. Grain is the smallest distinction we make in a set of observations.[9] Extent is the size of the widest thing that can fit into a set of measurements. Ecology has been particularly helped by new technology that allows wide extent while retaining fine grain. The study of landscape ecology of a modern sort was not possible until remote sensing, sometimes from hundreds of miles up in the sky. There is a security issue with oceanic satellite images with pixels smaller than 30 meters on a side because at 10 meters resolution, you can distinguish water from water with a submarine in it. The extent is the scope of all the measurements that were made or would have been pertinent had they been made. You cannot see something in an image if its signal is wider than the scope of the data.

The scaling tools are part of deciding levels of observation.[10] These are in contrast to levels of organization, which are explicitly not scalar. The latter are defined not by their size but by organizational characteristics of the things at those levels. Levels of organization are the familiar levels that define the different sorts of ecological subdisciplines like organismal ecology. Scalar levels and levels of organization are likely to be at odds.

And all of this is going on with an act as simple as seeing and recognizing a bee. It may be smaller and in the context of the flower it pollinates. But with a different sort of level, it can be seen as an organism at the organism level that may be in a population. Population is also a level of organization, but a level that is not necessarily bigger than the organism level. The reason for that apparent contradiction is that population is not just a collection of organisms; it is a collection of organisms that are required to be in some way equivalent. You, the reader, as an organism, are not in the population of mites that feed on flakes that come off your skin. They are called dust mites and are everywhere in most houses, and are responsible for some people having allergies to dust. Thus, you as an organism are physically bigger than the whole population of mites. Population is not necessarily higher than organism, unless the organisms are equivalent. The two organisms in figure 7 (a flea and an elephant) are so different that they cannot be aggregated in a population no matter how proximate they are. Populations have a distinctive relationship to energetics. Populations do not so much get bigger through birth.[11] In a population births only present vessels that may or may not be filled, depending on the resources available.

So it is important to distinguish level of observation from level of organization because the same thing can be in two entirely different sorts of level without any contradiction. Scientists have to know what they are talking about and how they are talking about it. It is only through the formalism described earlier that casual natural history can become the science of ecology. It was only after Beaufort that impressions of the wind at sea became formalized. So what seems like the proverbial walk in the woods can become scientific once formality is introduced. In this way, scientists can handle the cacophony of ecological players. There is almost always a rich and textured set of things and happenings, even in a simple ecological study.

The joy of models is that they are internally consistent. The burden of models is that they have to be consistent. Rosen notices that formal models are scale independent.[12] They use scaling equations as laws to identify relationships that apply across scales. Beaufort addressed an expression of the laws of aerodynamics. Notice that the same laws of fluid motion apply to large and small flying objects linked through an exponential equation. An example would be the equation for drag. Drag depends on the fluid medium through which the object passes. It also depends on the size, shape, and speed of the object. One way to express these relationships is by means of the equations for drag. There is no need to spell them out here, but drag relates to: (1) density of the fluid, (2) speed relative to the fluid, (3) cross-sectional area, and (4) drag coefficient, a number without units.

Intuitively, we know that as something moves faster through air, the air applies increased resistance. On a graph with air speed on the horizontal axis (the abscissa) and the drag plotted on the other axis (the ordinate), drag increases on an ever-steepening curve (figure 4*A*). But some objects, like flags, deform to catch the wind more, so drag increases faster than expectations for a rigid form. Other objects, like trees, deform so as to streamline. At first, leaves flutter to catch the wind like a flag, but then the tree bends and so streamlines to avoid the worst of the drag at times when the tree is in danger of toppling (figure 4*B*). Drag still goes up, but not as much as it would were the tree rigid.[13]

The power of the exponent in the drag equation is significant, which is why a gale of 40 miles per hour buffets you, but a hurricane of 70 miles per hour throws you about, and may even turn over vehicles. Here is Beaufort again. First, there is a cottage with smoke going straight up. Then leaves rustle and branches bend. In no time, force 8 for a gale is tearing off branches. By hurricane force 12, trees are uprooted and buildings are damaged. And all of this is because of the exponent on the drag equation. Double the speed of the wind and you get much more than simply twice the wind damage. The laws of aerodynamics are captured in a whole set of equations that give the rules for relating things of different size in different wind. The laws of aerodynamics are therefore scale independent. A paper airplane and a DC-10 both yield to those laws (figure 5).

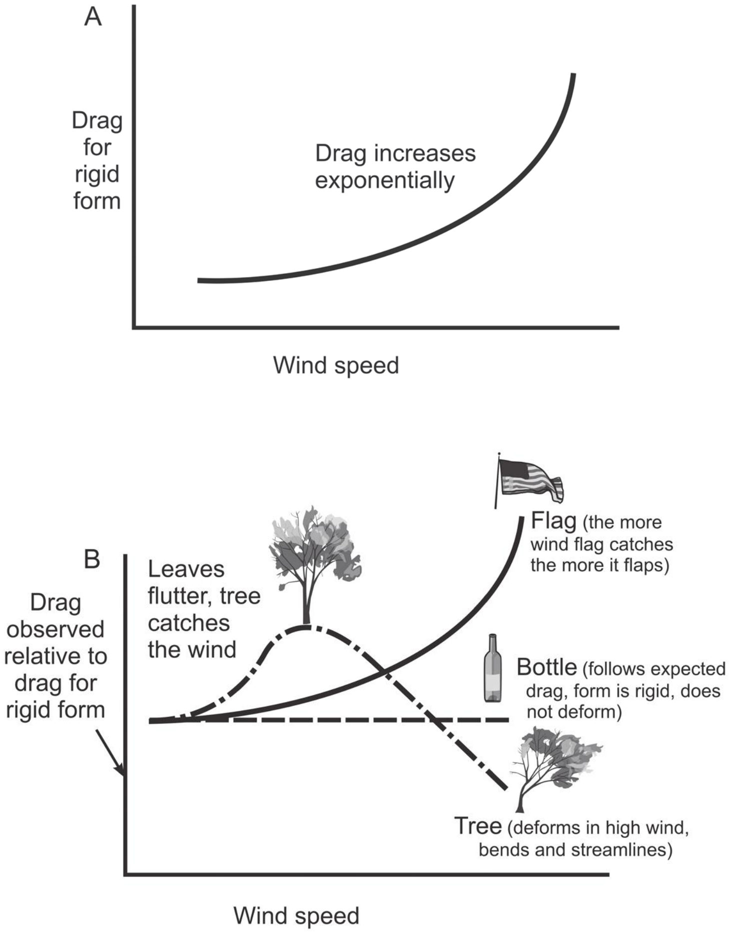

FIGURE 4. *A*. The exponential increase on wind speed against drag for a given rigid form. Graph (*B*) shows a dimensionless number, which is drag as expected on *A* for a rigid form divided by observed drag on the structure. Drag (observed) divided by drag (expected) cancels out the units to give a dimensionless number. Rigid structures like bottles do not deform and so measured drag over expected drag is a unit value. But flags flutter and keep catching the wind more as they flap harder. Trees have fluttering leaves and so in a light breeze they catch the wind more, like a flag. But then in high winds, trees bend away from the wind and so streamline, giving drags ever lower than expected, even as raw drag increases.

You do not want size to be specified explicitly in a model because then other sizes will not fit the model, even if the same principles prevail. Small and big things both fly, so we need a model that works for both big and small flying things. Ecology often makes the mistake of including size in models. We do not wish to anticipate the biome chapter in advance, but suffice it to say that the notion of biome, which is a model, is conventionally seen to apply only to big things, like the prairie biome that covers the Great Plains. In fact, the biome way of looking at

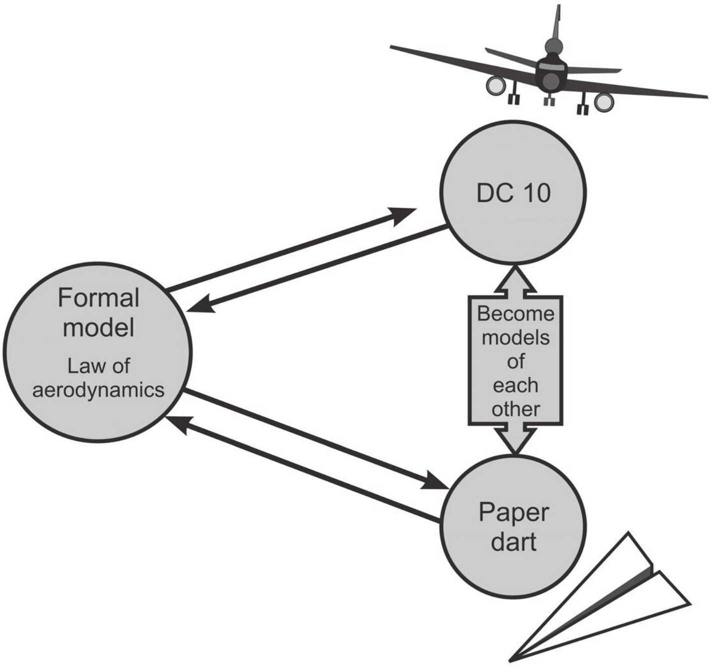

FIGURE 5. Rosen's (1991) modeling relation showing how a formal model of the equations of aerodynamics can be decoded into different flying objects. In turn, the material systems may be encoded back into the formal model. When that can be done, the two material systems become analog models of each other by compressing them down to only what the two have in common. The model for the child's toy is the DC-10 itself. The toy is also a model for the DC-10, so the analogy goes both ways. The analogy is not coded into words or equations; it is simply taken as a material equivalence. But the formal model *is* coded in symbols.

things also applies profitably to small things, like frost pockets. Frost pockets are an acre or so, but they are recognized in the same way we recognize big biomes. Conventional definitions of biomes would unnecessarily exclude frost pockets, even though the same principles apply.

The thing itself is not coded like the formal model, it just is. Translating material observables into the model is called encoding; you check to see how a material thing fits the coded specifications in the model. Yes, paper darts do greatly increase their drag if you throw them hard. Now it gets exciting! If you can encode and decode two material observables with a single formal model, something wonderful happens. The two material systems become equivalent so that you can experiment on one of them and use the results to predict behavior of the other. The principle is analogy, and all experimentation depends on that relationship. The formal model is a metaphor for the material system; it is a description, a representation. The relationship between the two material systems in an analogy is a compression down to only what the two things have in common. You do not refer to paper in the analogy between a paper airplane and the DC-10 because the DC-10 is not made of paper. Following the characterization of common ecological criteria in chapters 1 to 7, we devote chapter 8 to an in-depth explanation of the narrative and model relationship as a basis for a metatheory of ecology and the unification of ecology for use in management and research.

In ecology, there is a need for a framework that the scientist can use to organize experience; that is the challenge of ecology. We have just gone through the narrative description and basis for this book. Now, let us stand back to see the model and its rules that we propose for a unification of ecology.

This book erects that framework. Given the central role of scale in using formal models, the framework turns on scale. The concept is rich, requiring this whole book for a complete accounting of scale in ecology. However, at this early stage, we need to introduce briefly what we mean by scale, so that the word can pass from jargon to working vocabulary. Scale pertains to size in both time and space; size is a matter of measurement, so scale does not exist independent of the scientists' measuring scheme. Something is large-scale if perceiving it requires observations over relatively long periods of time, across large parcels of space, or both. With all else equal, the more heterogeneous something is, the larger its scale. An example here might be in comparing vegetated tracts of the same size. The vegetation that has more types of plants in more varied microhabitats more evenly represented is larger scale. It is more inclusive. Not only do things that behave slowly generally occupy larger spaces but they are also more inclusive and therefore heterogeneous. There are nuances of size all over the place.

In several topic areas of the ecological literature there is confusion because of opposite meanings between vernacular and technical terms. Evenly spread means low variation between locales. Ecologists use quadrats, square areas that are laid down for purposes of sampling vegetation. We often count plants in quadrats. Evenly dispersed vegetation leaves no sample area without hits because all the space is evenly covered. There will be no high values with lots of hits because a hit on one plant comes with a local surrounding space with no individuals in it; there are no clumps with which to score big. A frequency distribution is a statistical device where we rank order the values in samples. We note how frequently we encounter a given score. So with quadrat samples of evenly spaced vegetation where plants are counted in sample areas, there are no values of zero and no high values either. The variance of a frequency distribution is a measure of statistical dispersion that is a general measure of how wide the span of the numbers is in the samples. So, evenly spaced plants show low variance in the sample. That is whence the "underdispersion" comes from between samples in vegetation that is evenly spread out. With clumped plants, the sampling quadrats tend to hit clumps or miss them. So there are lots of zero samples and many full samples too. The

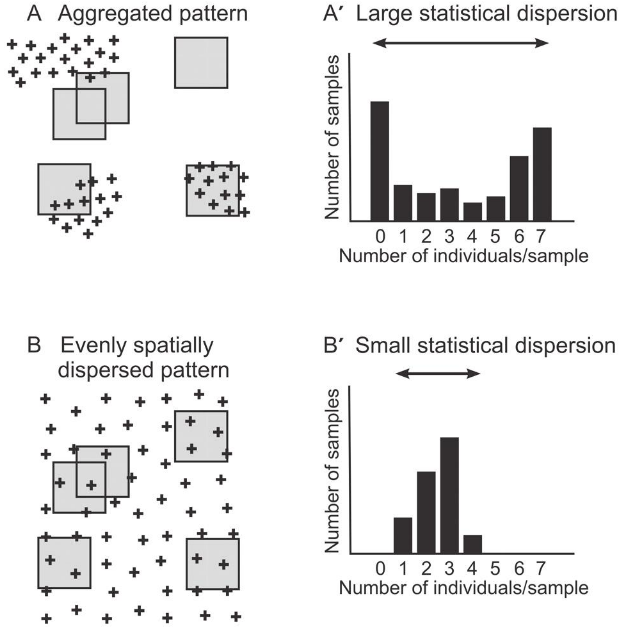

FIGURE 6. Two patterns of spatial distribution and their corresponding frequency distributions. *A*. Sampling clumped individuals with a quadrat gives mostly either empty or full samples, but with few in between. The variability of such a collection of samples is large, giving an overdispersed frequency distribution. *B*. With populations that are evenly distributed, all samples contain about the same number of individuals with little variation; the frequency distribution is accordingly underdispersed. Quadrats *A* and *B* are the patterns on the ground, while *A′* and *B′* are the respective statistical distributions.

frequency distribution shows a wide span of frequencies of zero to many, and so clumped vegetation is counterintuitively called overdispersed. The difference is that vernacular meanings refer directly to the thing or behavior, whereas the technical meaning refers to how one views the situation from some standard device or unit (figure 6).

So it is with “scale.” Something that is big, we call large-scale because it is large in and of itself. That is the way we couple the words *large* and *scale* throughout this book. Accordingly, we say that small things are small-scale. However, cartographers reading this book, and anyone else using their terminology, will be tearing their hair out. Our choice is either to ignore the sensitivities of a group of specialists or use *scale* in their counterintuitive way. We choose the former, but for clarity need to present the geographer’s point of view. A small-scale map in geography indicates that a unit measure like a mile will be very small on the map, so that a large area is represented. Conversely, a large-scale map shows small things on the ground with clarity, because a large-scale map makes them large. Therefore, a large-scale map must be of a relatively small area. A map of the entire globe would

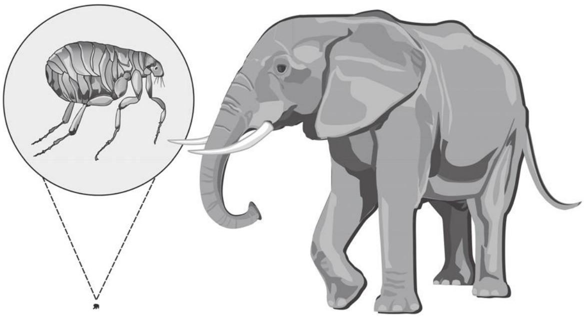

FIGURE 7. Both fleas and elephants are organisms, but their different sizes demand observation from very different distances. Note that the organismal form is very different with change of size, even though there is a head, a body, and legs in both cases. Those differences relate to scale.

be on the order of 1:50,000,000, a geographer's small-scale map of what we in this book call a large-scale structure. The technical meaning to a cartographer refers to the smallness of the one in relation to the fifty million. The vernacular meaning, the one we use throughout this book, indicates that the fifty million is a big number and the whole world is a big place; it is large-scale.

Ecology includes material and processes ranging from the physiology and genetics of small organisms to carbon balance in the entire biosphere (figure 7). At all scales, there are many ways to study the material systems of ecology. Let us emphasize that the physical size of the system in time and space does not prescribe the pertinent conceptual devices associated with different ways of studying ecology, although it may indicate the pertinent tools for observing, like remote sensing for big landscapes. Each set of devices or points of view embodies a different set of relationships. One ecologist might choose to emphasize physiological considerations while another might look at relationships that make an organism part of a population. But a physiological point of view is not necessarily small-scale. Elephant physiology includes more matter than does a whole population of nematode worms, and it is much bigger than the entire community or ecosystem in a small tidal pool or pothole (figure 8). The physiological differences in photosynthetic mechanisms between grasses define entire biomes in the dry western United States, so physiologists can think as big as almost any sort of ecologist. Brian Chabot and Harold Mooney have published an entire book on the physiological ecology of communities and biomes.[14]

The levels of organization refer to types of ecology. The types of ecological system are often ranked in textbooks: biosphere, biome, landscape, ecosystem, community, population, organism, and cell. We do not find that ranking useful, and call it the conventional biological or ecological hierarchy; each level therein we call

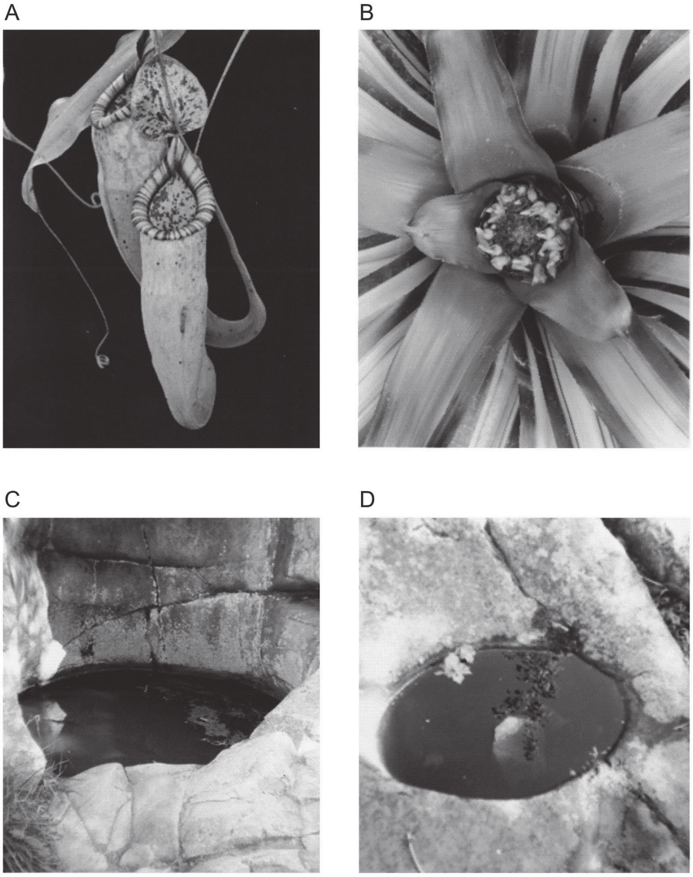

FIGURE 8. Pockets of water much smaller than some large organisms represent fully functional, self-contained ecosystems. *A*. The leaves of pitcher plants that trap insects. *B*. The ponds in the middle of epiphytes (photo courtesy of C. Lipke). *C* and *D*. Pothole ecosystems, from inches to meters across, made by boulders trapped in eddies in glacial outwash of the St. Croix River, Minnesota (photos courtesy of T. Allen).

a conventional level of organization. When seeking mechanisms, it is certainly a mistake to assume that explanatory subsystems must come from lower down the conventional ranking of levels of organization (figure 9).

In the literature of biological levels of organization, there are some branched variants of the conventional scheme, but the simple hierarchy captures the prevailing paradigm for grand, unifying designs for biology. Although many ecologists view themselves as working at a level of organization in the grand hierarchy, the levels defining the different types of ecologists do not strictly depend on the scale used by the respective scientists. The ordered relationships between the conventional levels of biological organization in fact offer relatively few explanations for the configurations that we seem to find in nature.

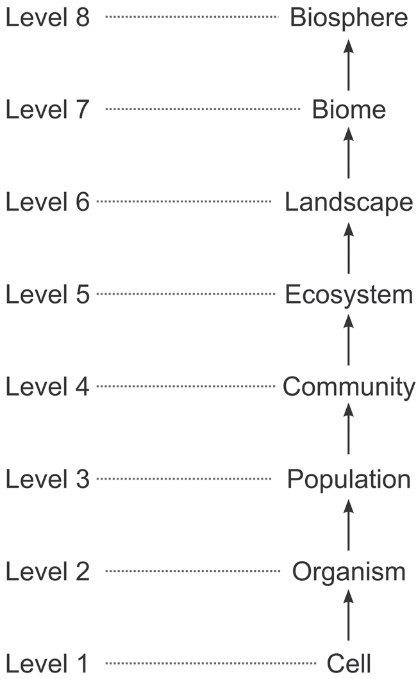

FIGURE 9. The conventional hierarchy of levels of organization from cell to biosphere.

Hoekstra personally experienced the logical inconsistency within the hierarchy of conventional levels of organization. The experience occurred on a warm, sunny fall afternoon in a beautiful remnant old-growth hardwood forest community, a factor that may well have contributed to the situation being still imprinted in his mind. It was a class field trip to collect plant community data when Hoekstra was an undergraduate student in Alton A. Lindsay's plant ecology class at Purdue University in the early 1960s. The conventional hierarchy puts ecosystems higher than communities because ecosystems are more inclusive and fold in the physical environment. Hoekstra was unable to get resolution to the inconsistency of how an ecosystem could be defined by a log on the forest floor that was at a smaller size than the forest community within which it existed. The conundrum was finally resolved through his work with Allen in the 1980s, and the application of hierarchy theory to provide the effective framework needed to differentiate between scale-defined levels of ecological criteria and system type. Their work on levels of organization distinct from levels of observation in the context of hierarchy theory was able to provide the explanation. This experience was part of Hoekstra's motivation for the research that led to the eventual development of this book.

If the ordering of conventional levels is often unhelpful, where can ecologists find powerful explanations for what they observe? The ordered sequence from cell to biosphere receives lip service as a grand scheme, but it is not the driver of ecological research activity. The conceptual devices that ecologists actually use in practice invoke explicitly scaled structures, not the generalized entities from the conventional hierarchy. The reason is that the conventional hierarchy is not scale-based, although its users think it is. We emphasize that conceptions invoked by conventional levels of organization are very important for ecological understanding, particularly when each is given autonomy separate from the grand conventional scheme.

We call the levels of the conventional biological hierarchy “criteria for observation,” or just “criteria,” to distinguish them from scale-defined levels. Criteria, as we use them, are not scalar but rather announce how one plans to study a slice of ecological material cut out for research. Criteria are the basis upon which one makes a decision as to what relationships are important in an ecological observation. The principal criteria in this book are: organism, population, community, landscape, ecosystem, biome, and biosphere. However, we do not use them as ordered levels per se, except to explain aptly certain relationships. We do not order the conventional levels by scale. When we do order by scale, it is within each one criterion. Not to anticipate the population criterion as it is considered in the population chapter under the population criterion, but suffice it to say that we look at pronghorn antelope as small populations of a doe and her fawns, as herd groups (which are not held together by relatedness), as a population in a favorable region, and finally as a large population that is the whole species.

crieria:
to single
out
certain
relations

We prefer a scheme that explains what ecologists find in practice, even if it lacks the feigned intellectual tidiness of the conventional hierarchy. Rather than a grand ordering scheme, we see the criteria as the prevailing means whereby ecologists categorize themselves. Almost all ecologists would put themselves in one or another of the criteria of the conventional hierarchy. Allen calls himself mostly a plant community ecologist. Hoekstra would answer to the appellation of ecosystem and landscape ecologist, with an applied bent of a wildlife ecologist. In a conversation about lakes Jim Kitchell, the distinguished limnologist, always says, “The zooplankton did it.” Kitchell noted that when faced with the self-same situation, Allen always starts looking for big patterns with the statistical tools of the community ecologist. “Tim, you always do that. I’ve seen you.” And it is true. Each type of ecology appears to involve its own style of investigation that follows naturally from the critical characteristics of the preferred conception. The contrast between those stylistic differences gives the relationships between the parts of this book. We dissect out the effect of scale within the type of ecology at hand.

Many general texts on ecology start with either organisms or physiology and build up to larger systems, extending to large ecosystems or the biosphere. Other texts start with the biggest ecological systems and work down through a process of disaggregation, looking for mechanisms and explanatory principles going down the hierarchy. They both make a fundamental mistake of scale ordering unscaled entities, but beyond that, they both invite a misplaced emphasis even if the scaling was adept. By choosing either order, up to biosphere or down to organisms and cells, the ecologist can easily be led away from considering interlevel relationships in the other respective direction. For an adequate understanding leading to robust prediction, you have to consider at least three levels at once: (1) the level in question; (2) the level below that gives mechanisms; and (3) the level above that gives context, role, or significance.[15] The conventional scheme also does not invite linking across to several levels away in the conventional scheme. It does not invite linking pathways at a subcellular level to differences between biomes or even the functioning of the whole biosphere, where indeed reasonable causality can be asserted.

In chapter 8, we give a formal account of linking levels in principle. But at this point, suffice it to say that there is subtlety and richness not found in the conventional scheme. Given our reservations for either a top-down or bottom-up approach separately, we feel that a different organization is appropriate, one that works with types of systems as alternative conceptual devices with equal status. We do not feel compelled to deal with a sequence of levels ordered by definitions that emphasize degrees of inclusiveness, although we use inclusivity when it is pertinent. The conventional ecological hierarchy is not wrong, but it is far too particular to serve as a framework for an undertaking as broad as we have in mind, a unified ecology. In unifying ecology we need to juxtapose levels and structures that come from distant places in the conventional order.

This book offers a cohesive intellectual framework for ecology. We show how to link the various parts of ecology into a natural whole. The prevailing lack of unity that we address comes from ecologists resorting to telling stories about special cases instead of rigorously defining the general condition. We like stories so much that we devote chapter 8 to them, but wish to go further than natural historical stories. There are too many tangibly different cases in ecological subject matter for us to retreat into a description of all the differences that come to mind. We offer an emphasis on what is similar across ecology, so order can emerge from a wide-ranging pattern. We take care not to be blinkered by the conventional hierarchy. When it fits, we use part of it temporarily.

The body of ideas we use has been gaining credence in ecology for more than the last quarter of a century under the rubrics of hierarchy theory, patch dynamics, scale questions, general systems theory, multiple stable points, surprise, chaos, catastrophe, complexity, and self-organizing systems. Although computers are not always used in the application of all these ideas, the mind-set that they have in common came from computer-based modeling of complexity. These collective conceptions are sufficiently mature for us to pull them together, with some new material, into a cohesive theory for ecological systems in general. The first edition of this book is still up to date, but the places where ecology has moved forward are now included.

The principles we use are those of hierarchy: a formal approach to the relationship between upper-level control that limits the outcomes from lower-level possibilities. Hierarchies of hegemony in unfair ranked societies are such a special case that we are not bothered by the political correctness sometimes stirred up by the word. Hierarchy is too important and generally useful for it to be sidelined for political reasons. The observer always has a scale of perception and a level of analysis that deals with the system as a complex observable. Hierarchy theory is not a set of esoteric speculations about ontological reality. It provides a hard-nosed protocol for observing complexity without confusion; it is an epistemology.

There are two separate aspects to observation. In hierarchies, content and context together generate significant behavior at each scale-defined level. We find the level by using a certain scale of observation. However, at a given scale, it is possible to recognize many different types of things. Which types are recognized and which are ignored come from the observer's decisions about what is to be considered important. "Criteria for observation" is the name we give to whatever it is that makes something important enough to be recognized in an observation or set of observations. Our hierarchical framework thus focuses on scale, on the one hand, and criteria for observation on the other. Criteria for observation identify what connections make the whole that appears in the foreground. Different things can be in the foreground at a given scale according to decisions about what to recognize. We get a much clearer view of both scaling effects and the consequences of defining a criterion in a particular way if we emphasize the difference between scale and criterion. The conventional hierarchy is not so careful, and therefore often leads to a muddle.

In ecology, the criteria for observation give rich perceptions, above and beyond the fact that many scales of perception are necessary to do justice to ecological material. For example, the organism can be conceived in many ways: the consequences of a genome and a vessel for housing the genome; a collection of internal regulated processes; an input/output system showing irritability; a system in a loop of action on the world, and a response to it; a structural mechanical system with scaling problems; and so on. This book investigates the richness of character of the objects of study that define the principal subdisciplines in ecology. We unite ecologists by the common strategies for observation that each group has developed parallel to, but separate from, their colleagues in other ecological subdisciplines.

As we apply hierarchy theory to ecology, we pursue relationships of a functional sort. In and around the material that ecology studies there is a tangle of flows of material and energy. We organize our treatment of ecology around those fluxes and the connections they embody. It is thus the unexpected limits on material flows that make ecology more than complicated physics and chemistry. This book looks at the material system, identifying fresh conceptions to account for the limits that appear to be ecological in nature. We attend to physical limits only occasionally and as necessary.

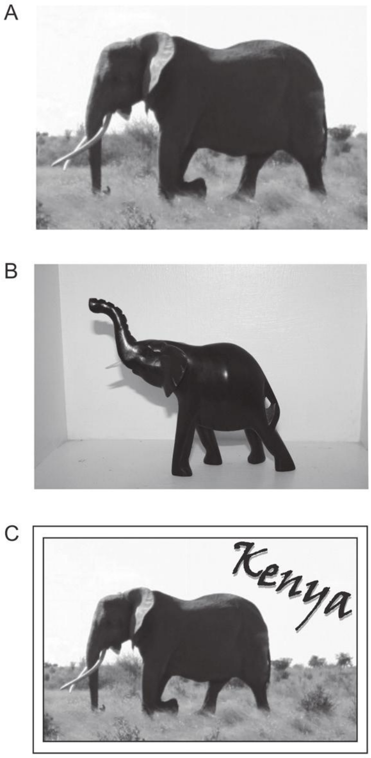

FIGURE 10. Three ways to realize the essence of a bull elephant. Nothing is scaled until it is realized. Notice that the realization in a photograph does not tell which of the other realizations is in the picture. Realizations are generally anchored to a purpose. *A*. The size of an individual free ranging bull elephant in East Africa (photo courtesy of T. Hoekstra). *B*. The size of a bull elephant sculpture. *C*. The size of a postcard about wildlife in Kenya.

We are careful to identify purpose in both the investigator and in the things we model. Organisms model purposefully; they have a model and tell stories. Purpose is distinctly not a physical thing. Behind the observables we see in biology are essences. The things we see are realizations of essences. A scale can be imposed only after there is a realization. Essence as a notion is unpacked fully in chapter 8. Here, we suggest essences of a particular bull elephant in East Africa (figure 10). Realizations are for a purpose. There are several realizations of the bull elephant essence: one offers experience during wildlife fieldwork; another is a sculpture for an office shelf. Yet another realization is a postcard of the bull elephant. The bull elephant essence can be realized as an object weighing several tons or only a few ounces. If you take the picture of the realization in the right way, you cannot tell which realization it is. Species and demes have coded purpose derived from evolution by natural selection. The purpose is transferred to the realizations, which then know and signify things beyond their experience. The biologist and social scientist needs to keep careful tabs on information.

A minimal set of premises underlies the organization of this book. They are:

1. Ecological processes and types of ecological structures are multiscaled. Each particular structure relates to a particular scale used to observe it such that, at that scale of perception, the entity appears cohesive, explicable, and predictable relative to some question the scientist asks. The scale of a process becomes fixed only once the associated scaled structures are prescribed and set in their scaled context. Scaling is done by the observer; it is not a matter of nature independent of observation.
2. The structures that match human scales of unaided perception are the most well known and are the most frequently discussed. Ecological processes are usually couched in terms of those familiar structures. The scale of those processes is prescribed by (a) those tangible structures and (b) a context that is also scaled so as to be readily discernable. A common error is to leave out discussion of the context. Meaning comes from the context, so meaningful discussion must include it. The principles derivable from observing tangibles deserve to be applied to the unfamiliar and the intangible.
3. At some scales of perception, phenomena are simpler than at others; the tangible system thus indicates powerful scales of perception. Predictability is improved if the scale suggested by immediate observables of the system anchors the investigation. The attributes of the particular type of ecological system distinguish foreground from background or the whole from its context. The criteria that distinguish foreground from background are usually independent of scale. It is therefore sensible to determine the appropriate scale of perception separately from choosing the type of system.

There is an observer in the system; only by knowing the location and activities of the observer can we avoid self-deception and start to make ecology a predictive science. In the end, ecologists are addressing the same challenges as did Beaufort. A complicated set of experiences need to be systematized and calibrated so that different observers have some frame of reference for prediction and communication. In this introduction, we have spelled out what Beaufort did so powerfully and eloquently, and have moved it over into the science of ecology. So the unlikely starting point of Beaufort and his scale now looks to be a more suitable place to start this book.

1

THE PRINCIPLES OF ECOLOGICAL INTEGRATION

Some scientific disciplines study objects distant from commonplace human experience. The stars are literally far away and the quarks might as well be, for all the direct experience humans have with them. By contrast, ecology studies a bundle of rich and direct human encounters. Most ecologists have fond memories of some childhood place or activity that not only stimulated a first interest in field biology but now determines what in particular they study. While limnologists might remember a pond behind the house and their first microscope, oceanographers may sit in their offices smelling imaginary sea air and wandering along sea cliffs of summers long ago. Henry Horn has communicated one such story about himself:

> Several blocks from my childhood home in Augusta, Georgia, the street lost its pavement and continued as a dirt track through an abandoned field that was growing to woodland. I was not supposed to go there, especially alone, but I did. What drew me were birds, trees, and butterflies. The butterflies collected in prodigious numbers at the edges of evaporating mud puddles. In my memory they were mostly Tiger Swallowtails (*Papilioglaucus*) and Zebra Swallowtails (*Eurytides marcellus*) and the large Cloudless Sulphur (*Phoebis sennae*), and they sat in little groups according to their kind, all pointing in the same direction. At the time, I wanted to learn more about birds and trees, but I just wanted to enjoy the butterflies, creeping close and lying on my belly very carefully to avoid the stains of red clay that would betray where I had been.[1]

Horn says that as a graduate student, he camped by the Potholes Reservoir of eastern Washington. He was washing camp dishes by a stream between North and South Teal Lakes, with water striders all over. Ever the kid, he sheepishly admits that his curiosity led him to put some detergent on the surface of the pond, which of course wrecked the surface tension on which the striders were striding, and they sank. Surface tension now held the skaters in the water. Filled with remorse, he managed to rescue all but one, patting them off, rinsing them, and air drying them. On release, they strode away. Horn is a particularly curious ecologist. Tony Ives claims, "I spent huge amounts of time when I was a kid outdoors—my Dad was the director of the Mountain Research Station at CU [University of Colorado]. But I was a pretty lousy boy naturalist. Maybe that's why I'm a theoretical ecologist!"[2] It takes all sorts. Ecology is a very "hands-on" study, where the scientist often goes to natural places and looks at other living things in their own habitats.

The things we study in ecology seem very real. Nevertheless, ecology is a science and is therefore about observation and measurement more than about nature independent of observation. It is easy for a physicist tracking subatomic particles in a bubble chamber to remember that science must work through observation and has no direct access to ontological reality. Science in its practice is not about truth and reality; it is about organizing experience and predictive power. In the end, it is about developing compelling narratives. Compelling narratives lead to commensurate experience. Prediction does not so much tell us that we are right about what is happening as it makes our narratives more convincing, so the listener starts to see things the same way as the narrator. Then they can talk and deepen each other's understanding. Almost everything a physicist measures comes through the filter of some gadget. For the ecologist, it is harder to remember that measurement is not reality, for out in the field where birds sing and flowers bloom, all the human senses are flooded with experiences that have an ecological basis, if we would only do the ecology. Even so, ecology is a matter of organizing and challenging perception, with reality always at least one step behind the screen.

Our bottom line is that there may be a real external world out there with things in it, so we are far from committed to antirealism. But there is much that lies between our explanations of our experiences and external physical reality. After the science is all over, many will prefer to believe that what science finds gets ever closer to external reality. Nobody can ever prove that, but such a belief is just fine. It is part of the myth, a belief of modernist science. Since we put narratives front and center in our discussion, we have to respect myths. The myth of science is what keeps some scientists dealing with the tedium of collecting endless data. And in ecology, the data collecting might even be uncomfortable, as when it is raining. Peter Greig-Smith taught Allen ecology, and Greig-Smith had never been known to cancel a field trip because of rain. And it rained a lot in North Wales. Yes, the myth keeps some of us going; it raises science to a metaphysical discourse. So what is our objection to realism in ecology if we are not committed to there being no external reality?

Realism is just fine when the work is over. In retrospect, the consistency of results and the predictions of experiments can reasonably encourage a belief that science has progressed closer to reality. What exactly that means is less clear. As George Box said, "Essentially all models are wrong, but some are useful."[3] Box is emphasizing that in the act of representation, experience passes through a set of symbols that bear an arbitrary relationship to that which is represented. There is no true representation in the way that there can be better or worse approximations. Representation passes through arbitrary symbols, and so is never in terms of that which is represented. Approximation is always in the same terms as that which is approximated. Hourglass, water clock, grandfather clock, and quartz crystal atomic clocks successively keep better time. Critically, it is time all the way, so degrees of being correct are easily understandable. Analog recording devices, such as a long-playing record, have physical connection from sound to groove, and groove back to speaker. In a sense, the record approximates the sound in reproduction through the speaker. But a CD or DVD is different. Digital media introduce arbitrary symbols, for instance a certain number for middle C. In that way there is the critical step to representation that is never in the same terms as the original. Representation can be better or worse for certain stated purposes, but then purpose is not a material thing so is not something that can be approximated. Representations are not approximations and that throws a monkey wrench into the notion of scientific models closer to reality. Better models that serve a purpose is something else, which is captured in Box's "but some models are useful."

The difference between analog and digital is the same difference as between a sign and a signal or symbol. When Sherlock Holmes sees footprints on the ground outside a window, he is seeing marks, signs that someone was there. The sign can be elaborate, as in the pattern of the detail of the sole and its wear. There is no signal in the footprint, it is just the mark left by someone standing. When Holmes develops an account of what happened, and ties the sole of the footprint to the shoe of a suspect, something is added. It has interpretation from the observer. The interpretation is a model, and all models have symbols with meaning, and that is more than just a sign. When Abraham Wald looked at planes returning from air raids in World War II, he saw the places where the planes had been hit.[4] The orthodox view of the remedy was to look at the pattern and reinforce the planes where they had been hit most often. That model suggests that planes tend not to be hit in the places where there were no bullet holes. Wald said to reinforce the planes where there were no bullet holes. His logic was that planes that get hit there never come back. Orthodoxy mistook bullet holes for the reality of planes getting hit. Wald suspended judgment for holes until he had worked out that the holes he saw were places that planes could survive. Data are only data, and are always in context.

Meaning is in the context, and that is a separate part of modeling that involves abstraction. There is also abstraction in the orthodox interpretation, but in taking the data as reality, interpretation is blunted. One can make assertions as to what was really going on with the hits, but they are best suspended until the context is interpreted.

Realism during the conduct of investigation gets badly underfoot while the science is being done. Notice how Holmes is flexible in his interpretation. He is open to the signs implying any of a large number of scenarios. If the terminology is expected to reflect reality, it is not surprising that terms are held with an unreasonable stubbornness that blinds insight. Realism encourages inflexibility. Thus, realism leads to semantic arguments (squabbles over terminology) and so wastes a lot of time. Just because species seem real, there is no reason to suppose that they are real where other taxonomic levels such as genus or family are taken as only abstraction. In the end it is all abstraction, so asserting the real is an unfaithful anchor. Ernst Mayr noted a remarkable coincidence between the numbers of recognized scientific species of birds in New Guinea with the number of bird types recognized by locals not trained in ornithology. "But I discovered that the very same aggregations or groupings of individuals that the trained zoologist called separate species were called species by the New Guinea natives," Dr. Mayr said. "I collected 137 species of birds. The natives had 136 names for these birds—they confused only two of them. The coincidence of what Western scientists called species and what the natives called species was so total that I realized the species was a very real thing in nature."[5] Had he chosen some other group, such as lizards, the coincidence would not have been there. The issue is that birds to humans are like computer clones to the IBM personal computer. Birds are different on the inside, but they share our favored input and output devices, as do computer clones. Humans and birds address the world orally and perceive the world visually. So of course we recognize differences between birds that birds do. Their breeding patterns depend on things we readily recognize. Lizards live in a world where they taste the atmosphere. We are no good at that. Species are not only an abstraction, they are an anthropomorphic abstraction. They represent the degree of variation that humans can readily grasp. For a taxonomist, species have a gestalt. Genera and taxonomic families are simply useful abstractions that are based on criteria different from those taxonomists use on species. All are abstractions. There may be something in reality going on with regard to a particular species in a particular study, but it is still wise to delay asserting reality as long as possible so as to avoid jumping to conclusions. The authors are antirealist only to that point.

If ecology had avoided realism, as Arthur G. Tansley argued before his time,[6] we might not have wasted the twentieth century trying to find out what is a community, really. The Clements/Gleason debate (discussed in chapter 4) is largely semantics driven by different levels of analysis connected by one word, "community." Tansley met with many thoughtful social scientists, and even Sigmund Freud. When everyone else was arguing whether climax vegetation is a real state, Tansley went into a very sophisticated caveat:

> One last point, we must always be aware of hypostasizing [reifying] abstractions, that is, giving them an unreal substance, for it is one of the most dangerous and widespread vices through the whole range of philosophical and scientific thought. I mean we must always remain alive to the fact that our scientific concepts are obtained by "abstracting from the continuum of sense experience," to use philosophical jargon, that is by *selecting* certain sets of phenomena from the continuum and putting them together to form a concept which we use as an apparatus to formulate and synthesize thought. This we must continually do, for it is the only way in which we can think, in which science can proceed. What we should not do is treat the concepts so formed as if they represented entities which we could deal with as we should deal, for example, with persons, instead of being, as they are, mere thought apparatuses of strictly limited, though of essential value.* Thus a plant community is an essential concept for purposes of the study of vegetation, but is, on the other hand, an aggregation of individual plants which we choose to consider an entity, because we are able to recognize certain uniformities of vegetational structure and behavior within the aggregation by doing so. A climax community is a particular aggregation which lasts, in its main features, and is not replaced by another, for a certain length of time; it is indispensable as a conception, but viewed from another standpoint it is a mere aggregation of plants on some of whose qualities as an aggregation we find it useful to insist . . . But we must never deceive ourselves into believing that they are anything but abstractions which we make for our own use, partial synthesis of partial validity, never covering *all* the phenomena, but always capable of improvement and modification, preeminently useful because they direct our attention to the means of discovering connections we should otherwise have missed, and thus enable us to penetrate more deeply in the web of natural causation.
>
> *Footnote [of Tansley]: A good example of the hypostatization of an abstraction, exceedingly common 40 years ago, but now happily rare, is the treatment of the process of natural selection as if it were an active sort of *deux ex machina* which always and everywhere modified species and created new ones, as a breeder might do with conscious design. (Tansley 1926, 685–86, emphasis and quotation marks in original)

One manifestation is the attempt by some organismal ecologists to justify assertions that concrete things like organisms are real and material, whereas communities and ecosystems are mere abstractions. Allen, once challenged in this way, was expected to assert: "No, communities are real." He did not, but rather preferred to

FIGURE 1.1. Portuguese man-of-war.

argue that organisms are abstractions. Some “organisms” stretch the definition of organism, as when a Portuguese man-of-war (figure 1.1) possesses not one but three genetic lines in its body, one for the sail, one for the body, and a third for the tentacles.

Some argue that a fungal strain found all over the Upper Peninsula of Michigan is a single organism because all the disconnected threads have the same genome

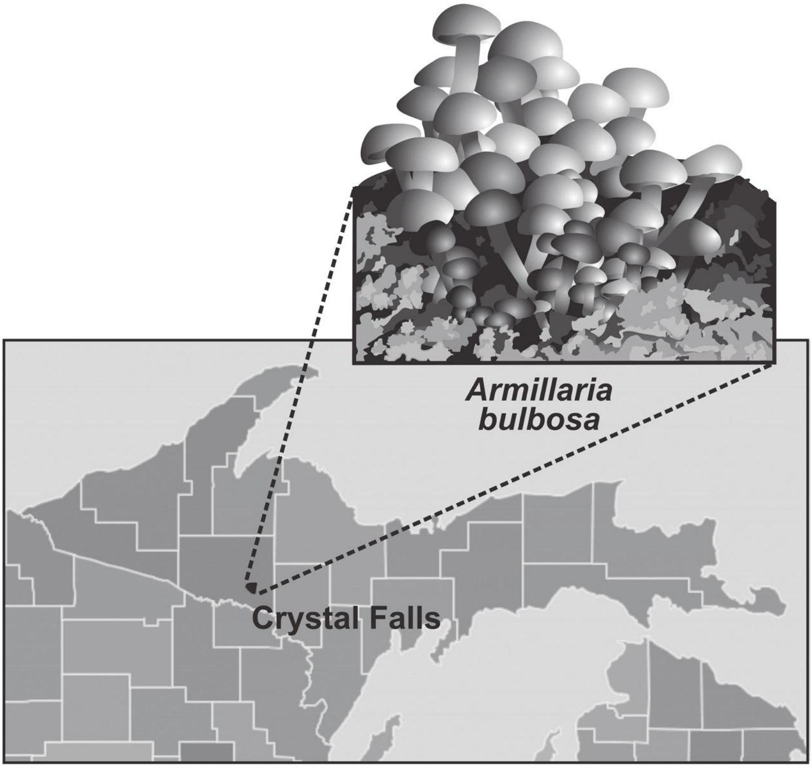

FIGURE 1.2. *Armillaria bulbosa* found in the vicinity of Crystal Falls in the Upper Peninsula of Michigan.

(figure 1.2).[7] Organisms are a meaningful, significant conception in ecology, but significance is a human assertion, not a material issue.[8]

We do not want data for their own sake; we want them for the more general condition that lies behind the data that we got on a given day. Data for their own sake are a scientific version of stamp collecting. Let us drive that point home by introducing neural nets.[9] They are used, among other things, to calculate whether someone applying for a loan will actually pay it off. The data consist of aspects of the loan seeker reported in the loan application. Some things are obviously connected to the reliability of the prospective borrower, such as income and number of dependents.

Other things are less obviously relevant, like whether the borrower owns a house and how many years the applicant has lived in it. Even numbers like postal zip code might be included. All those bits of information about the applicant are put in a row of information that is fed into the top of the net.

Often, the top row is connected to the net, with each item connected to all the nodes in one level below the row of data. The input to the nodes is determined by a weight as to importance of the input. That is randomly assigned in the beginning, but the connections in the net are reweighted in response to success in a known training set of data. The training set is fed in repeatedly in random order. The inputs to a node are summed in some way, and the sum is fed into a function that says in the end whether the node should fire (like a nerve fires) a signal down to the next lower level of nodes. In the case of a prospective borrower, the net fires on down to give a value for a single node at the bottom of the net. If that value is high, the loan would be granted, but if it is low, the loan is denied. There is a data set coming from people who applied for a loan where the banker knows that the borrower did or did not pay it back. If the net gets it right, the net reinforces the previous adjustment in weights. If the net gets it wrong—say, denying a loan to someone who actually did pay it back—the previous adjustment in weights is reversed. After a lot of training, the net gets good at telling reliable applicants from others in the training set (figure 1.3).

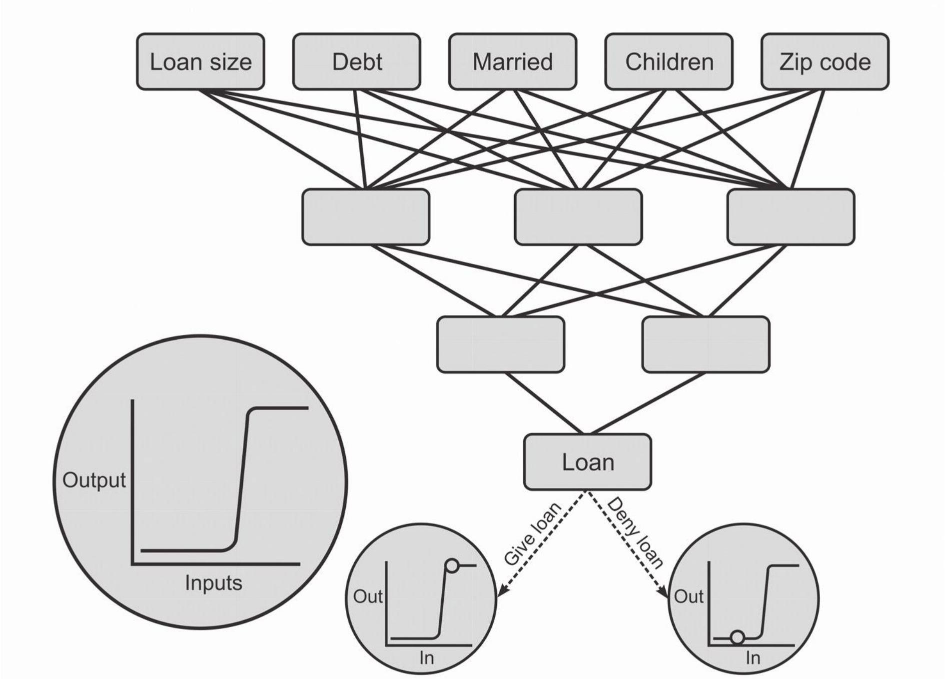

FIGURE 1.3. Neural nets compress a vector of inputs in a loan application to predict if the applicant will pay off the loan. The training input data sets include who paid on past loans. Output of each node (box) is fed into the node below, with all inputs summed in the node below. The sum is fed into a function. If the input sum is high, the function tells the node to fire. If the input sum is low, the node does not fire an input to the node below. The connecting lines are weighted according to the ultimate success in predicting who paid. Success = adjust weights in same direction. Failure = reverse last time's adjustment. Weights change over time. When weights are found to be accurate predictors, the net can be used to predict the credit rating of new applicants. The usual conversion function is inset here, but others are allowed. Incomplete connections are also allowed in the net, and some may even jump levels to give privileged information to lower nodes.

When the net is calibrated to be good at predicting, it is then used on people who are applying for a loan, to determine whether they will pay it back. If you have ever applied for a mortgage, your information went through this process. In the training process, there are two data sets presented to the net. One is the training set described earlier, to which the net responds; the other is the test data set. The test data are real too; they include people for whom it is also known whether they paid or did not. Critically, the net is not allowed to respond to the test set, only the training set, and so the net forgets any experience of the test set. The net is not allowed to remember what it was told in the test, and it is not allowed to adjust the weights on the nodes in response to the test set. Memory of the training set is embodied in the adjusting weights. As a result, it is possible to track the improving accuracy on the training set, as well as the accuracy on the test set, which the net cannot remember it has seen. Early on in the training process, if prediction on the training set improves, so do the predictions on the test set.

And now we get to the reason why we have discussed neural nets, how they relate to why we collect data in the first place, and why we rely on them in science. As the training set predictions get very good, suddenly the predictions on the test data get worse (figure 1.4). The reason is that the last adjustment in the face of the training set was an improvement because the net was starting to memorize

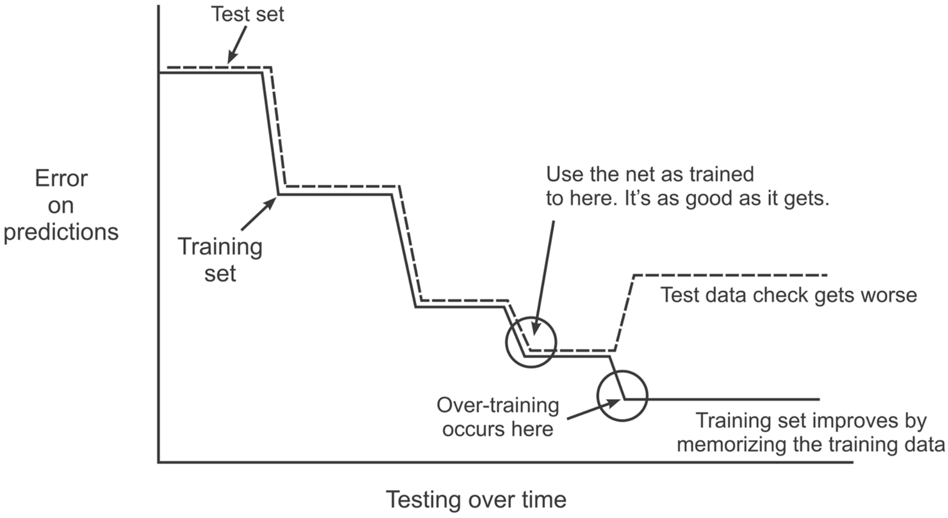

FIGURE 1.4. Results of training neural nets over time improves prediction in both the training set and in the test set, to which the net is not allowed to respond and so cannot remember. The test set is independent. Improvements occur in both sets over time with training; until, in the end, the training set improves while the test set results get worse. At that point, the net has been overtrained and is memorizing the particulars of the training data, not the generality the data represent. Use the results that the net achieved just before overtraining. It is the best you can do.

the details of the training data. It was coming to the conclusion that this person is a reliable borrower because the net registered that person not as a representative data point, but as a data point it had in particular seen before. In data there are two signals. One is the general pattern that the data represent; that is the signal you want to detect and calibrate. The other signal is the data that you collected in particular on the day you sampled. If you had collected data on some other occasion in exactly the same way you collected the first data set, the new data would be a bit different. You will have sampled different if equivalent people. And the new data would continue to be different on every new occasion you collected more data, even though all those new data sets would still reflect the general phenomenon. Those differences between successive samples of the same universe are incidental and reflect that any data set has its particulars that do not matter. Science does not want data for their own sake; investigators do want data as they reflect the general condition the scientist is investigating.

Ecology may be dealing with a fair reflection of what is behind the veil of our observations, but we have no way to know. As scientists, we deal only with observations, observables, and their implications. We do not rely on assertions that any ecological entity is real in an ultimate sense. We try not to be biased in favor of observables in tune with unaided human perception. However, we do acknowledge that understanding in ecology, as in all science, involves an accommodation between measurements and models that are couched in distinctly human terms.

THE OBSERVER IN THE SYSTEM

If a physicist studying quantum mechanics chooses to suppress the role of the observer in the system, the consequence is wrong predictions. This is called the observation problem, the dilemma of reliance on observation to gain insight into the world that is above and beyond the specifics of the observation. Biologists also have their observation problem, but they prefer, for the most part, to postpone dealing with it. In this volume, we tackle the observation problem head-on. We discuss ecology driven by observation, even if we appear at times nuanced and abstract. The subtleties on which we insist actually take using data more seriously than does the mainstream realism. Data are based on the observer's paradigm as much as externalities. At this juncture, ecology needs a generally acceptable body of theory, for which we propose a theory of observation that ecologists can use to acquire data. If ecology is to become more predictive, it will have to be more careful in recognizing the implications of its observation protocols. Without the ecological observer, there can be no study of ecology. Even at the grossest level of decision making, when the ecologist chooses what to study, that act influences the outcome of the investigation. When one chooses to study shrews, there is an implicit decision not to study everything else. In that implicit decision, most other things ecological, such as trees, rivers, or ants, are excluded from the data.

Ecological Phenomena and Definitions

Phenomena in ecology, as in science in general, are manifestations of change; there can be no phenomenon if everything is constant. Robert Rosen said in *Anticipatory Systems*:

> The observing procedure is the very essence of abstraction. Indeed, no theory, and no understanding, is possible when only a single mode of observation (i.e. a single meter) is available. With only a single meter, there cannot be any science at all; science can only begin when there are at least two meters available, which give rise to two descriptions that may be compared with one another. (2012:214)

Recognizing those changes that constitute a phenomenon must be preceded by observer decisions about what constitutes structure. To accommodate change, there must be some defined structure, a thing to change state. Our observations involve arbitrary structural decisions, many of which revolve around making or choosing definitions. Definitions are not right or wrong, but some give us more leverage against nature's secrets than others. For example, scientific species names or their vernacular counterparts seem to be powerful ways to categorize living things. Nevertheless, the things included under those names are as arbitrary as any other named set. This manifests itself in disagreement between taxonomists about distinctions between species. A misidentified plant involves using the wrong definition; while the species mistakenly used may be a good species, it was the wrong one on that occasion. Right and wrong is a nuanced distinction.

One might argue that species are abstractions and that their very abstractness is the source of their arbitrariness. However, even something as tangible as a tree is, in fact, arbitrary. The entire army of Alexander the Great camped under a single banyan tree. The army was large, but the tree was old and had grown, as is the nature of the species, by sending roots from its limbs down to the ground. At first the roots are threadlike, but after a long time they thicken to become tree trunks in their own right. Thus, one tree can become a forest. Contrast this with a clone of aspen. As the stem establishing the clone becomes large enough to spare reserves, it spreads out its roots. These long roots periodically send a branch above the ground that then grows into an apparently separate new plant. The organic connections between the trees in a clone are quite as strong as the limbs connecting the trunks of an old banyan tree. The only difference between a clone of aspen and a banyan tree is our perception. We do not see underground. An earthworm's eye view might see the

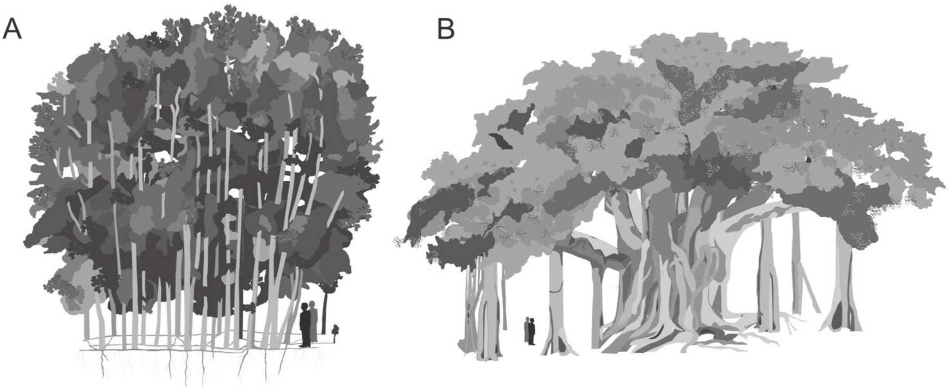

FIGURE 1.5. *A*. Aspens spread their roots, and new above-ground stems come from those roots. But we cannot see the connection and so call an aspen clone a set of separate trees, even though the clone is as connected as for the single banyan. *B*. Banyan trees come from one tree stock, but roots drop down to make new tree trunks. We can see the connections and realize that it is one tree. Allen and Hoekstra have been introduced into both figures as small people for scale.

banyan as a grove of separate trees but the aspens as all one organism. Thus, even what constitutes a single tree is a matter of arbitrary human judgment (figure 1.5).

Definitions are generally based on discontinuities that have been experienced or are at least conjectured. In the case of the banyan tree and the aspen clone, there are two critical discontinuities. One is the separation between tree trunks; the other is the separation of whole banyan groves or whole aspen clones. Note that the separation between the tree trunks in both banyans and aspens is only a matter of degree. That is why a formal definition is so important, because without it there is ambiguity as to whether the tree trunks or the whole interconnected collection of tree trunks constitute the organism. Note that neither the separate trunks nor the collection of them all are truly the proper level of aggregation to assign to the class “organism.” However, a given discourse about aspens or banyan trees has to be consistent in the meaning of the words it employs. Simply make a decision and take responsibility for it.

A definition is a formal description of a discontinuity that makes it easy to assign subsequent experience to the definition. How steep a gradient of change will have to be to be assigned the status of a “discontinuity” is a matter of decision. The observer experiences the world and decides whether the experience fits the definition. With a new definition, some experiences that were once within the defined class are now excluded, while others may be added. By some definitions, a banana plant is a tree; by other definitions, even a banana plant ten meters high would not be a tree. If “tree” is defined as plant material above a certain height, bananas can be trees. The “trunk” of a banana plant is mostly fleshy, sheathing leaf bases, not woody stem (figure 1.6). The stem arises later and grows up the middle of the

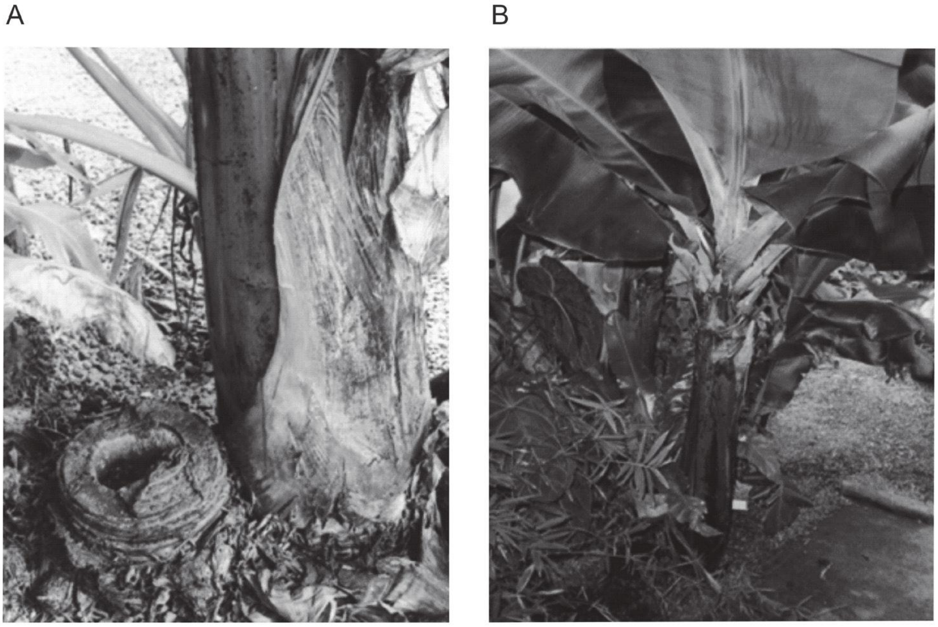

FIGURE 1.6. A cross-section of a banana tree identifies that the "trunk" is almost entirely fleshy sheathing, not woody stem. (Photos courtesy of C. Lipke.)

sheathing leaf bases when it is time to make bananas. Therefore, a definition of tree that insisted on woody stems could leave even the tallest nonflowering banana out of the class "tree." In a sense, a banana is a tall herb, but the distinction has a human origin; nature does not care what we call it. Behind all acts of naming are implicit definitions. For all we know, nature itself is continuous, but to describe change, we must use definitions to slice the world into sectors. The world either fits into our definitions or not. Either way, all definitions are human devices, not parts of nature independent of human activity.

Definitions, naming, and identifying critical change are not the only arbitrariness in scientific observation. The observer uses a filter to engage the world. The filter chosen by the observer is as much a matter of human decision as is the definition of structures. Because we cannot measure anything in infinite detail, differences too small for the instrument to detect (fine grain) are filtered out of the data. In this case, the smallest differences fall below the level of resolution of the study. Furthermore, any difference that takes too long a time or is too large spatially to fit into the entire sweep of the data (extent) will not survive the filter of the data-collecting protocol. Beyond this, any signal that the instrument cannot detect will be missed. For example, a light meter will not measure pH. Sometimes one hopes for a surrogate signal, as occurred when Carol Wessman used remotely sensed radiation as a measure of lignin concentration in forest canopy, which in turn is a correlate of soil nitrogen.[10] All these aspects of the scientist's input filters are arbitrarily chosen, and they all influence what the ecologist experiences.

Reduction has been a useful tool in that explanations often derive from taking things apart so as to address the parts separately. Reductionism takes reduction one step further into a philosophical realm. It asserts that the whole can be known from its parts. Reductionism takes the question at hand to be self-evidently interesting. It also says that the scientist knows that the chosen level of reduction is somehow the right one. For instance, cancer can be explained and investigated to the point of fighting cancer if one reduces the issue to cell division and failure of moderating signals across cell membranes. Reductionism gets the scientist quickly into data collection. Actually, cancer hormones controlling blood supply appear to be an equally useful explanatory device. Critically holistic approaches also use reduction, but the question asked is not taken as self-evidently interesting. The holist does not assert the level of reduction at the outset. The holistic question says, "I do not know what I need, but it would probably have this and that property. Has anyone seen one?" Ecology needs that level of openness and flexibility.

It is a mistake to presume that studies of small things, like biochemistry, are reductionist. It is just that the holism had been done earlier and well, so the assertion of the level of explanation is well founded. For instance, Hans Krebs took advantage of the many details worked out by other researchers in his discovery of the Krebs cycle.[11] His breakthrough was less in his effort put into the details of *What* and was more in his bigger holistic question of *Why*. Why did the tissue continue longer in its capacity for oxidation when he added citrate? Ah, it closes the loop, an answer that applies to a higher level of analysis, not a lower level of detail. Significance is always a higher-level property. We now know that biochemical cycles like the Krebs cycle are everywhere, and why we should expect them (see chapter 3, figure 3.5*A*).

EMERGENCE IN BIOSOCIAL SYSTEMS

The difficulty with reduction is that the emergent properties of the whole cannot be derived from the parts or predicted unless they have been seen before. It is not possible to go into water with a device small enough to sort out individual water molecules, and then pull out a wet one. Wetness is a property of water in aggregate. In a universe only the width of a proton, there can be no chemistry. That only occurs in a larger universe. In recent years, much has been made of emergence in complexity science, but at least some of it is far too grand. Emergence need not be a big deal.

Emergence is always some version of finding something new. The trivial case is when you have looked at something in a new way, perhaps more precisely, and found something you did not know about. Sometimes emergence arises because of a wider universe of discourse, as in the proton example, where a wider universe gives chemistry. Less trivial are aspects of emergence that are seen to arise out of positive feedbacks. Positive feedbacks require some sort of gradient to drive the expansion of system components. At the top of the gradient, matter or energy exists concentrated in an unlikely arrangement. In biosocial systems, the gradient starts in the environment. As matter energy is taken in from the top of the environmental gradient, perhaps as food, a gradient is set up inside the biosocial system. The bottom of that gradient is manifested in effluent arising from degrading the inputs. The material or energy at the bottom of the gradient is in a much more likely arrangement, having arrived there by randomizing processes. The randomization follows from the second law of thermodynamics under which things run down. The total gradient from environment back to effluent runs down. But inside the biosocial system, the gradient is fueled by inputs and therefore remains stoked up, far from equilibrium. The internal gradient does not run down unless death occurs. Gradients are needed to drive the positive feedbacks underlying emergence (some positive feedbacks show inexorable decrease, but that is a matter of how the system is specified—a gradient is still needed to drive values down). *A* and *B* increasing under positive feedback do not do so free; it takes an expenditure of energy to drive the increase or decrease. That energy comes from using the external gradient. Physics talks of the second law in closed systems running down, but living systems run up, not down. Quite the opposite of denying the second law of thermodynamics by becoming more organized, life uses the second law for organization.[12] Emergence has been described as some sort of explosion, but that misses a critical part of it. The emergent property is a constant that appears when a racing positive feedback encounters some limit that can usually be cast as a self-correcting negative feedback. An emergent property does not pertain until the explosion encounters some sort of limit.

As we come to the end of the industrial age, there remain many holdovers from the industrial posture. Industrial problems at the cutting edge in Victorian Britain were met by building larger structures and applying more power. Solutions like that have always cost more with increase in size. Meanwhile, problem solving in the postindustrial age involves smaller things that offer immediately cheaper solutions. Telephones and computers are only slightly reduced in price, but each generation comes with greater speed and effectiveness. The price of engines with moving parts generally goes up as fast as inflation because they use industrial principles. A big holdover from the industrial age is ambitious space travel, which should be seen as a bigger engine going further. The idea of colonies on any other celestial body is simply an outdated Industrial Revolution posture at best. Putting up human-made satellites around our own world is a different matter because the issue there is not moving things a long way, but is rather about compressing information about our whole planet, a postindustrial notion. Rockets taking humans to Mars are really overgrown steam engines whose time has passed.

There are general principles here. The difficulties with industrial age ambitions are that they invite emergence. Emergence involves catastrophic failure of the system manifested before. Applying huge force to a larger engine offers a steep gradient that will set off a positive feedback. The crash of the space shuttle *Challenger* started with the failure of O-rings, which were a small part of the system. With that much power applied to takeoff, the gradient was plenty steep enough to cascade dysfunction upscale, which it did. Disaster frequently comes from a process of emergence that had never been seen before. The expensive correction is a matter of prediction, but by then the engineers have seen it before. We cannot predict emergence until we have seen it at least once. Big power and large systems will continue to rise in cost faster than humanity can pay for them. The problem is that emergence in principle is connected to steep gradients. The industrial age is slipping away because it is not a viable strategy in the face of emergent failure and its cost. The postindustrial problems are also a matter of emergence. They have manifested themselves in deep financial global retrenchment caused by steep gradients of information, not steepening gradients of energy and matter, as they were in the industrial age. Any steep gradient becomes an accident waiting to happen, albeit a gradient of information or matter/energy. A cluttered high shelf is an invitation for something to fall.

THE MINIMAL MODEL

When there is danger of over-reduction, the concept of the minimal model is helpful. The minimal model gives predictions from the smallest number of explanatory principles. Thus, the small number of explanatory principles in a minimal model comes from using the highest level that can be contained inside the system to be explained. Occam's razor invokes the principle of parsimony, where one should pursue the simplest explanation. Kevin Kelly has been working toward a deeper justification for Occam's razor.[13] He cites statistics that show those using Occam's razor retract results less often and sooner than those who do not use it. Recently, he has shown why we would expect simpler models to find something with the properties of truth more often. His analysis uses proposed truths, a troublesome notion, but we know what he means. He finds that Occam's razor will find these truths faster, with fewer steps, and more often than more complicated models and procedures.[14] One wants to find the highest logical type that is pertinent. Einstein urged us to make things as simple as possible, but no simpler. The razor cuts away everything else as superfluous.

The relationship between prediction and ultimate reality is not at all clear, so justifying cumbersome modeling by reference to ultimate reality should not be accepted as an excuse for ignoring minimal reduction of Occam's razor. Howard

Pattee suggests that reducing below the level of meaning lets the issue at hand slip through the investigator's fingers like sand.[15] The chemistry of the ink will not enlighten the reader as to the meaning of this text.

THE THEORETICAL BASIS FOR SCALING AND INTEGRATING ECOLOGY

Structure and Process

Up to this point we have emphasized structure and entities. However, another facet of levels and scale uses a process-oriented conception of entities and patterns. Biological systems are very much a matter of process. In fact, what appears to be distinctly structural in biology can often be seen as part of a process. For example, the human body consists of material that flushes through in about a seven-year cycle (even the bones). There is very little left of the "you" of seven years ago. In biology, often one process reinforces one or several others; in turn, the first process is reinforced by those it has influenced. That self-reinforcement leads to persistent configurations of processes. That persistence of process clusters explains how biological structure can appear concrete, although the substance of that structure is in constant flux.

Let us start with something tangible and identify its underlying processes before moving to more abstract biological structures. Solid, concrete things are surrounded by surfaces. The surface is all that we see of most things because it is the part through which the whole communicates with the rest of the universe. Although one might conceive of surfaces as passive and having nothing to do with dynamic processes, surfaces are places where the dynamic forces dominating the internal functions of an entity reach their functional limits. The skin of an organism corresponds with the furthest extent of the internal circulation system. The skin also coincides with the limits of many other fluxes.

Science looks for surfaces that define things with generality. We seek things relevant to systems according to many criteria, things that are detectable even when one looks at the system a different way. An entity that shows this persistence we call "robust to transformation." Processes held inside a stable surface usually reinforce each other. In social systems, language and commerce reinforce the limits in each other at an international boundary. Trade uses language and a common language facilitates trade. In the tree, growth puts leaves in the light, then photosynthesis provides material for growth; growth and light each refer to the limits on the other. These mutually reinforcing processes give a set of surfaces that can also be seen as mutually reinforcing, one surface for each process. For example, in lakes in summer, the surface between warm water above and cold water below is the surface at which oxygen and nutrient status coincide; the process of oxygen use

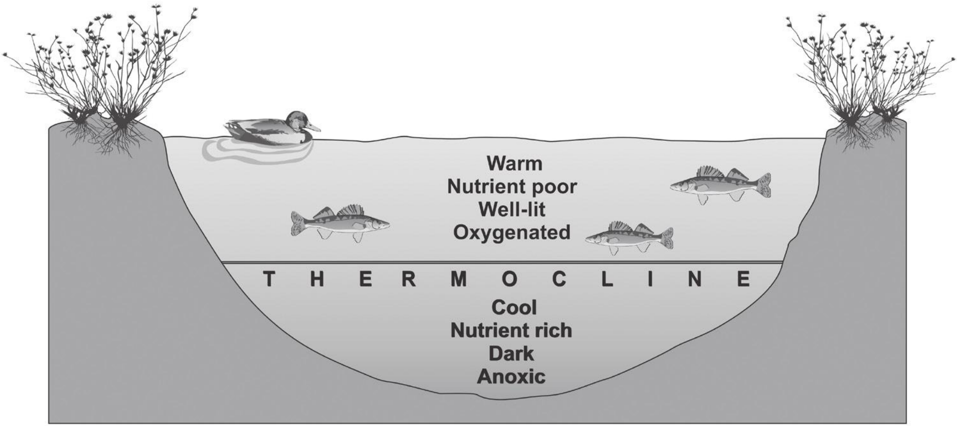

FIGURE 1.7. In northern temperate lakes in the summertime, warm water mixes above while cold water mixes below, separated at the thermocline where there is a very steep temperature gradient.

depletes nutrients, so there is high oxygen concentration above the thermocline and high nutrient status below (figure 1.7). The multiplicity of devices that can be used to detect the surface of an entity that is robust to transformation reflects the multiplicity of the processes associated with that surface. The thermocline is an emergent that then constrains processes in the vicinity.

Some plants grow and establish apparently new individuals by vegetative growth. A collection of individuals formed in that way has connections between genetically identical individuals. The aspen clones discussed earlier are a good example. The surface of a clone, as with many surfaces, is identifiable by sets of processes that press the surface outward. For example, the process of water transport within the clone is rapid. Internal processes reach the surface, but then attenuate rapidly. Fluxes associated with processes inside slow down at the surface; this can be used as a demonstration of the presence of a surface. For example, a radioactive tracer moves rapidly through a clone by virtue of the interconnections between individuals. The radioactivity moves rapidly up to the surface of the clone, but only very slowly across it into the outside world. Surfaces are places where signal is attenuated, stopped, or changed in some way.

We have come to understand that processes cascade upscale through fractal geometry (fractals get an extended treatment in chapter 2). The dominant processes can be recognized by the texture of the structures the processes leave behind. For instance, a naturalist recognizes the species of a tree by the texture and outline resulting from the growth processes for that species. That is how biologists know which species is before them without going into the technical details of fruits or leaf shape. The tall, narrow trees in French impressionist paintings are Lombardy poplars. Lombardy poplars are columnar, not bushy like most open-grown trees. Lombardy poplars are so common in a British or Continental European scene that the British Army has a simple classification of trees: fir trees, poplars, and bushy top tree.[16] We suppose that is enough information for aiming artillery, although a silviculturist would find it a bit austere. The poplars referenced there are just the variety called Lombardy, with other varieties of poplar being bushy top. Lombardy poplars have a mutated growing point that causes the branches to diverge more slowly than most poplars, giving a distinctive fractal dimension to that variety. They are reproduced by humans by taking cuttings, and so the mutation persists. The growing point cascades upscale in a fractal manner, leaving a characteristic form behind. Small events that involve stems branching and growing leaves eventually give large outlines. Behind the structure lies a particular set of processes.

As we have seen, surfaces disconnect the internal functioning of entities from the outside world. The disconnection is significant, but not complete. Therefore, the observer has to judge whether or not the disconnection is sufficient to warrant designating a surface. That judgment is what makes all surfaces arbitrary, even natural surfaces that are robust to transformation.

The inside of a natural entity is strongly interconnected, as in the case of the trees in the clone. The inside is relatively disconnected from the outside across the surface. The strong connections inside and the weak connections to the outside are a matter of relative rates of fluxes. This applies not only to tangibles but also to intangibles like functional ecosystems. Large terrestrial ecosystems are composed of the interaction and integration of plants, animals, soil, and climate. The surface of an ecosystem is not tangible, but is rather defined by the cycles of energy, water, nutrients, and carbon. If the ecosystem has the integrity to make it a worthy object of study, then it must have a stable surface in some terms that reflects that integrity. You generally cannot see an ecosystem edge with the naked eye, but you could detect it with the use of radioactive tracers and a Geiger counter. The surface of the ecosystem makes it an entity. Connections inside the cycles are strong relative to the connections to the outside world. Tracers put into an ecosystem will move around inside the ecosystem much faster than they move out of the ecosystem. The relative rates of movement define the ecosystem surface, even if the ecologist cannot see it literally.

Other intangible entities, like the regional cultures within the United States, are also a matter of relative movement. There are more social contacts within a cultural region than between the region and the outside world. Because of the separation by water, the islands off the Carolinas each had their own dialect, one switching the *W* for a *V*, as in some Dickensian accents (Sam Weller pronounced with a *V*). All of that was homogenized with the establishment of just one school to serve all the islands.[17] Even with the inroads of mass culture through television, the islands as a whole still maintain an accent that sounds more southwestern English than American. It is a maritime accent, like those in pirate movies. The reason is that the social contacts are strong locally but weak to the mainland. The ocean defines the cultural surface.

Similarly, ecological populations are held together and are recognizable less because one can see the whole population and more because of the mutually reinforcing processes that bind the individuals together. Where population members are on the ground is often not distinctive, but the breeding patterns are a force for unity and identity. The process of breeding involves finding mates in a shared habitat. The process of survival in a habitat is often heritable. The shared responses of individuals to habitat define and are defined by what is inherited.

The boundaries of organism are often distinct, the process that holds them together being a contained physiology. Biomes are often large and may have distinct boundaries. Allen was on a field trip to the inland forests of Washington State. His hosts took him through many miles of unbroken Douglas fir–dominated forests. Then they crossed a ridge, and all was different. A juniper scrubland presented itself. The boundary between these biomes was distinct, measured in meters, not kilometers. Allen's host, Tom Spies from the Andrews Forest site, told him that the juniper desert scrubland emergent on that ridge goes all across the great basin for almost a thousand miles. Physiology integrates organisms as blood circulates carrying energy and hormones. As to homogeneous vegetation across thousands of miles, fire and animals are the only things that move fast enough to integrate large biomes.

Higher and Lower Levels

A relationship between levels can be considered as the relationship between the internal functioning of an entity and the behavior of the whole. Lower levels are characterized by internal functioning, while the upper level relates to the whole entity. Many problems can be translated into the relationships between levels. Evolution by natural selection is an example of a concept that links levels. The lower level is occupied by the individuals that reproduce with varying amounts of success in between birth and death. A Darwinian view sees the aggregate of those successes and failures as determining the character of the upper level, the population.

We are now in a position to discuss the principles that govern the relationship between higher and lower levels. We recognize five interrelated criteria that order higher levels over lower levels.

1. *Bond Strength.* In the discussion of surfaces, we have emphasized that the connections inside an entity are stronger than connections across its surface. The surfaces that separate the entities of a lower level can be part of the bonding that unites those entities as parts of the upper level (figure 1.8*A*). Relative to the strong connections

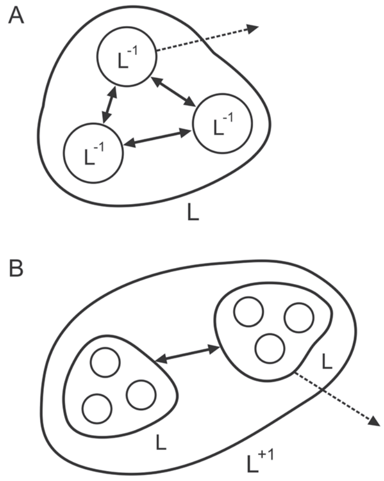

FIGURE 1.8. *A*. L is the level in question; L^{-1} is the next level down, and L^{+1} is the next level above. *B*. The weak connections of L^{-1} to the outside world beyond L become the strong connections with level L^{+1}.

within the parts, only weak signals pass out through their surfaces to make a link between the parts. The parts can only communicate with the relatively weak signals that pass out through their own surfaces. The weak connections between the lower-level entities, between the parts, become the strong connections that give integrity to the entity at the upper level, the whole. This explains the principle of weaker bond strength. The higher the level, the weaker is the strength of the bonds that hold entities at that level together (figure 1.8*B*). Sometimes it is possible to see the bond strength at a given level by breaking the bonds and measuring the energy released. Note that breaking chemical bonds in a fire or an explosion releases much less energy than breaking the bonds inside atoms in an atom bomb or a uranium fuel rod.

2. *Relative Frequency*. Behaviors that might appear to be directional can be seen as cyclical if the set of observations has a wider extent. For example, eating can be described with much more generality as an activity that is repeated at a certain frequency, rather than as a series of items consumed. In fact, describing all sorts of human activities as recursions allows a completely new set of insights. Herbert Simon once noted that if a resource is to be found everywhere, we replenish it often. Simon, an economist, sees this as a matter of inventory; only keep a short inventory of readily available resources, but keep a long inventory of scarce material whose supply is erratic. There is no point in planning too far ahead with something as universal as air, so we breathe at a relatively high frequency and keep a three-minute inventory of oxygen. Water is common but not ubiquitous, so we drink many times a day and keep an inventory of a few days. Our biological requirement for food has been set by the

hunt, and so we eat somewhere between once and five times a day. We have an inventory of food in our bodies that could last a few weeks, at least. Time to exhaust supply should be a good measure of past selective pressures. If there is total system shutdown, the greatest urgency of renewal applies to the shortest inventory.[18] Humans are quite hard to replace because of food to reach maturity and time to learn how to be effectively human. There is therefore an inventory on each one that is a generation long. Long-term human knowledge has an inventory on it of threescore and ten years. Levels of observation can be ordered by the frequency of the return time for the critical behavior of the entity in question at a given level of organization. Higher levels have a longer return time, that is, they behave at a lower frequency. Thus, in a coherent account of ecosystems, many nutrients cycle once a year, but the whole upper-level ecosystem accumulates nutrients over centuries and millennia.

3. *Context.* Low-frequency behavior of the upper level allows it to be the context of the lower level. The critical aspect of a context is that it either be spatially larger or more constant over time than the lower level for which it is the context. It is not always transparent what the critical constancies are that the context offers. Sometimes the context involves change, but that change always happens; the constancy is the always of “always happens.”[19] Seeds germinate based on a favorable growing season that has not yet happened but which is predictable from warming, moist conditions in the spring. Summer always follows spring, and that change is itself a constant on which temperate plants rely. If the environment is always changing, then what becomes dependable is the constancy of the change. Weedy plants thrive in disturbed habitats, and it is the constant upheaval that they find reliable and persistent.

4. *Containment.* Despite being unequivocally upper-level entities in the social hierarchy, dominant animals do not contain lower-level individuals. However, there is a special class of important systems where such containment is a requirement for existing at a high level. These systems are nested, where the upper level is composed of the lower levels. Organisms are profitably seen as nested in that they consist of cells, tissues, and organs. Of necessity, they also contain the parts of which they are made (figure 1.9*A*). In nested systems, the whole turns over more slowly than its parts, and the whole is clearly the context of its parts, so the nesting criterion aligns with the frequency and contextual criteria. However, the reverse is not true because the criteria of frequency and context do not depend on the system being spatially nested. For example, a reliable food supply behaves more slowly than the animals that depend on it, but it does not contain or consist of the animals it feeds. Unlike frequency and context, the bond strength criterion probably does apply reliably only to nested systems.

Nested systems are very robustly hierarchical in that the containment criterion corresponds to many other considerations. The nesting keeps the order of the levels constant even when the observer changes the rules for relationship between levels. Consider a hierarchy that goes from cells to communities. Plants may be seen as the synthesis of cellular interaction by physiological processes: cells to

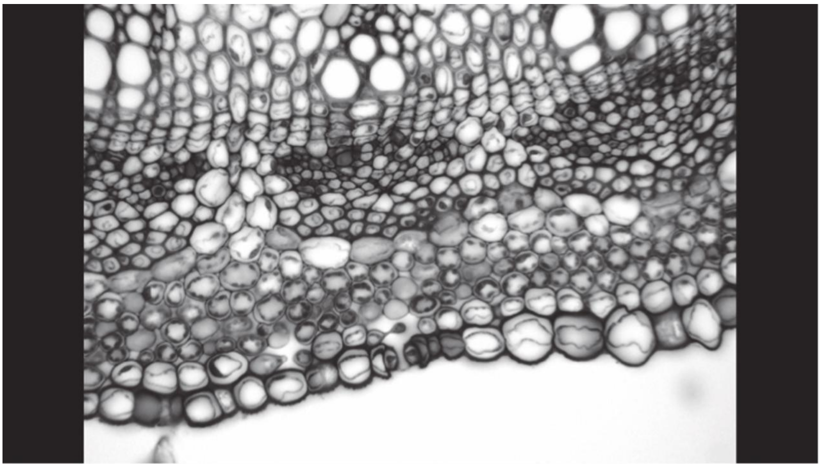

FIGURE 1.9*A*. Cells aggregated to form a tissue are an example of a lower-level being contained by a higher level. (Photo courtesy of University of Wisconsin Botany.)

plant on physiological rules. However, the whole plant can be seen as a part of the ecological community on rules of assembly that bear no simple relationship to the physiological processes that built the plant as an individual. The switch from physiology to species associations causes no confusion because the nesting keeps things straight. Western medicine usually casts the human body in nested terms. As a result, the cells interact through membranes and pores between cells. But the organism of a whole human is joined largely through processes of hydrodynamics. The pumping heart makes the kidneys work, which is why heart failure engenders kidney failure. At a higher level, humans interact socially through symbolic schemes of speech that have little to do with either cellular or whole body physiology. But it is all kept straight by the nesting of the human hierarchy. That the nesting is a choice rather than a necessity is evidenced by Eastern medicine often using lines of force and critical points rather than nested body parts.

By contrast, in non-nested systems, the containment criterion does not apply, and so it cannot be used to keep the system ordered; accordingly, non-nested systems have to rely on some other criterion, and they must use that one criterion from top to bottom of the hierarchy. A change in criterion in a non-nested system creates a new hierarchy that only incidentally shares certain entities at particular levels with the old hierarchy. Pecking orders or food chains are non-nested in that the top of the hierarchy does not contain the lower levels (figure 1.9*B*). The relationship between the plants and the grazer is the same as the relationship between the herbivore and carnivore; once again, higher levels do not contain lower levels. If one takes an individual in a food chain and then considers that same individual's place in the pecking order, this amounts to a switch to a new non-nested hierarchy

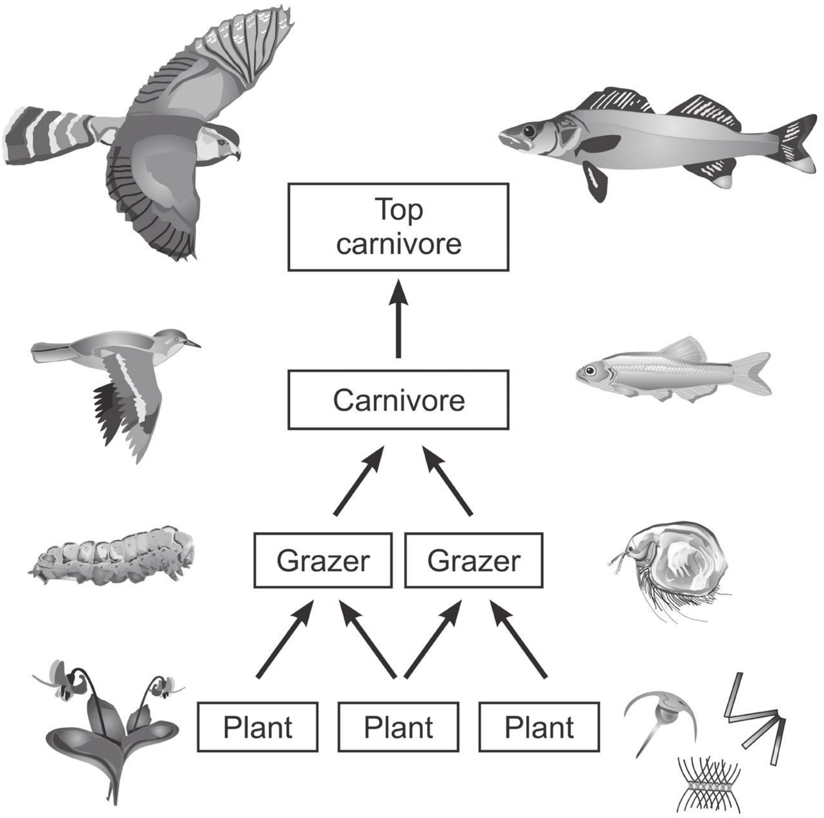

FIGURE 1.9*B*. Food chain hierarchies represent non-nested systems where the higher levels do not contain the lower levels.

that only incidentally articulates with the first non-nested hierarchy. From top to bottom of the food chain, the ordering criterion is always eat and be eaten; the social hierarchy of dogs at a kill is a different matter that needs separate consideration. The ordering in non-nested systems can easily be reversed, when "eats" is substituted by "depends on." A mapping of ungulate herds moving is slower and higher order than the movement of wolf packs that behave as satellite units. The wolf pack seen as moving around the herd as a lesser entity is working in the principle of "depends on" rather than "eating." Hierarchies must be used with precise criteria and for a particular purpose. Hierarchies are largely decisions about how to organize experience, not things independent of human concerns.

5. Constraint. For our purposes, frequency and constraint are the most important criteria for ordering levels. Upper levels constrain lower levels by behaving at a lower frequency. Constraint should not be seen as an active condition. Upper levels constrain lower levels by doing nothing or even refusing to act. For example, one should never underestimate the power of impregnable stupidity. High-frequency manipulations of elegant ideas can be held in the vice of persistent misapprehension. Contexts are generally unresponsive to the insistent efforts and communications of things held in the context. Constraint is always scaled to longer time frames than those used by that which is to be constrained. It is in this manner that deans constrain their university departments by doing nothing. A budget is announced perhaps once a year, and the departments must simply get used to it. Deans should not deal with every new circumstance, even those promoting worthy projects. Poor administrators change their minds when lobbied by their more powerful underlings. In such situations, there is nothing the unit as whole can bank on.

In this regard, consider an advancing dune system with plenty of sand still available for further advance and a suitable prevailing wind. Dune systems do not start easily from nothing; they mostly expand from established dunes that interfere with airflow. Dunes lose sand from their windward sides, and this sand is then deposited by the slack air on the leeward side. A flat terrain will allow the sand to blow away without forming self-perpetuating irregularities in airflow. If an intermittent river cuts the front of the dune system back, then that is a temporary limitation that is not a serious constraint. In time, the slow process of dune building will recoup the loss. However, if the river always returns before the dune system has had time to build past the riverbed, then the dune system is contained and constrained, and can never cross and establish a system on the riverbank on the far side. The individual floods are not the constraint; it is the "always" in the return of the river that makes the river a constraint. If the dune system could affect the course of the river, then there would not be constraint.[20] Note in this example that constraint can only be described at the right level of analysis.

Constraint is so important in the ordering of levels in ecology because it allows systems to be predictable. Predictability comes from the level in question being constrained by an envelope of permissible behavior. Predictions are made in the vicinity of those constraining limits. When a system is unpredictable, it has been posited in a form that does not involve reliable constraints. Any situation can be made to appear unpredictable, so predictability or otherwise is not a property of nature, it is a property of description. The name of the game in science is finding those helpful constraints that allow important predictions. Those are the models that George Box called "useful." Science would appear to be less about nature and more about finding adept descriptions.

In ecology, constraints are sometimes called limiting factors. An impossibly large number of factors could influence the growth of algae in a lake. Predicting population size is only possible when some critical known factor becomes constraining. Diatoms are microscopic plants that require silicon for their glass cell covering (figure 1.10).

Silicon is not very soluble in water, which is why we drink liquids from glass. But glass is soluble enough that soft museum specimens sitting in unopened jars

FIGURE 1.10. Diatoms develop a cell covering of glass. Accordingly, they are sometimes constrained by silicon supply in the water. (Photo courtesy of T. Allen.)

for a century come out just a bit crunchy from silicon of the container getting into the solution and precipitating on the specimen. If a lake has low silicon concentrations, then diatoms are reliably only a minor component of the plankton. Recent research on the great new dams in China indicates that the impoundments cause silicon to be removed from the water that flows through the dams because it is instead deposited at the bottom of the lake in the remains of dead diatoms.[21] The ecological result downstream and in the oceans and in the estuaries is a great depletion of diatoms, as well as the release of deleterious red tides in the estuary. Much ecological work is incidental recording of local considerations, but the new work on the effect of dams is an example of ecology that matters. When the constraint is lifted and silicon is abundant, diatoms may become abundant only if other factors are not limiting.

When control of a system changes, the situation becomes unpredictable. With a change in constraint, new factors take over the upper level. A good example is the switch in control that occurs in an epidemic of spruce budworm that can be seen as a switch in constraints. For decades, the budworms are held under the constant control of birds that eat them. Any increase in budworms will feed larger populations of birds that crop the budworm populations back down again (negative feedback) (figure 1.11). Budworms eat the young parts of shoots of conifer trees.

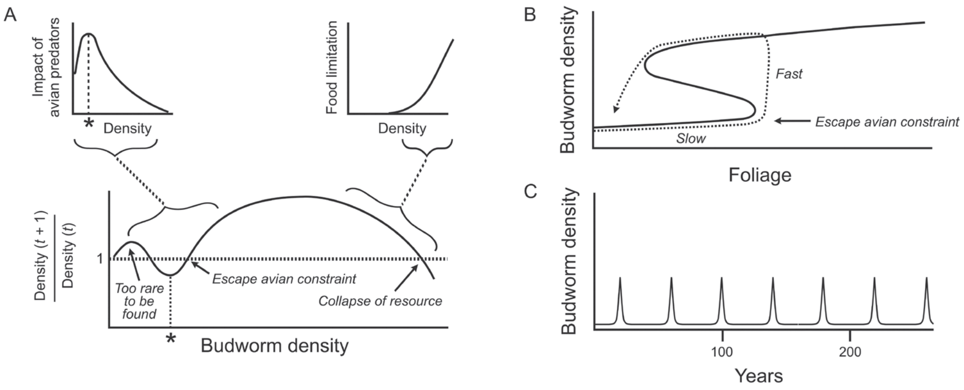

FIGURE 1.11. *A.* When the density in the next time unit is greater than the density in the present time unit, the tendency will be for the population to increase in numbers. That is the tendency of the system when it is above the unitary line. It will proceed to higher densities until there is a constraint to hold the density in the next time unit the same as the present. The unitary balance creeps slowly toward higher densities as host foliage increases in quantity. At a critical stage, the population can reach a high enough density so as to escape predation as it slips above the unitary line to grow to outbreak numbers. The population increases uncontrolled until the pest eats itself out of resources. *B.* The increased levels of pest population with increased foliage are gradual until there is a short, rapid outbreak of the pest that finally destroys the resource base. *C.* Plotted on a time axis, the population accordingly shows intermittent outbreaks approximately every thirty years (Holling 1986).

The avian constraint on the budworms allows the trees to grow. As the trees grow, the budworms have more food and can handle the losses to bird predation more easily. The budworms are still constrained by the birds, but their population creeps higher. Eventually the trees are sufficiently large that they support a population of budworms that is still relatively small, but is almost large enough to saturate bird constraint. Any bird can only eat so much, and bird density has upper limits controlled by factors like nesting sites or territorial behavior that have nothing to do with bird food. Sometime during a critical period in the constraint of budworms, when bird populations are at their maximum, a chance event like a pulse in worm immigration can increase the budworm population. Being limited by some other factor, the birds cannot respond as before by increasing their numbers (figure 1.11*B*). The effect is to increase the budworm population a small but critical amount where the budworms can grow faster than the birds can eat them, a critically different situation. Budworms increase in numbers leaving bird control behind. The constraint on the budworm population is broken as the birds lose control of the increasing population of insect larvae (figure 1.11*C*).

Since there has been a breakdown of the constraint, there are problems of predictability. When exactly the constraint will be broken is unpredictable, although it can be generally expected during certain time windows. The whole system is no longer under the reliable constraint of the birds, and it takes on the explosive dynamics of the epidemic. The well-behaved system under the bird constraint is of no help in predicting what happens when that constraint no longer applies. During a budworm explosion, a new constraint takes over, the growth rate of the uncontrolled budworm population. Soon, yet another constraint ousts the budworm growth curve as the limiting factor, when all the food disappears as all the mature trees are killed. Pest starvation constrains the system and it returns to low budworm population densities.

The constraint of the birds gives the trees time to grow; the system behaves slowly. When the constraint changes, the budworms become the upper constraining level for the system dynamics. The system suddenly starts behaving according to the fast dynamics of the budworms. From the vantage point of the old avian constraint, the new constraint comes as a complete surprise. A reordering of levels to give control to a new upper level gives unpredictable behavior.

The Conventional Framework: A Point of Departure

We now unpack our definitions that underlie the six criteria we use to order and unify ecology.

The Criterion for Organisms

We have already identified organisms as arbitrary (see introduction, figure 9). Botanists are more comfortable with this notion than vertebrate biologists because many plants regularly reproduce vegetatively and only gradually become autonomous and separate; however, vertebrates resemble ourselves in having an unambiguous physical boundary and obligate sexual reproduction in most species. Nevertheless, in humans, conjoined twins draw attention to the arbitrariness of even human organisms. Physical identity is an important point in the biology of placental mammals like humans because confusion between self and not self invites either spontaneous abortion or uncontrolled fetal parasitism. Both of these conditions reduce fitness.

Organisms are generally of a single genetic stock, although fruit tree grafting can lead to chimeras, branches where the outer layer of cells belongs to one partner of the graft while the core of the shoot belongs to the other. In botany, the formal definition of a new generation focuses on the reduction of the organism to a single cell, a fertilized egg, or a spore. The reduction to a single cell gives the genetic integrity of the individual. Even so, some population biologists view the parts of a plant, particularly grass shoots, as competing individuals. The shoots from a single root stock are called ramets,[22] as opposed to genets, which are plants derived from a single seed. Genets have genetic identity, while ramets have physiological integrity, as in a single stem. While the genetic and physical discreteness of organisms clearly fails in such conceptions, the individual grass shoot is adequately defined by its physiological autonomy.

Thus, to be an organism, a being should have at least one of the following: (1) genetic integrity, because it comes from a single germ line (egg or spore); (2) a discrete bodily form; and (3) physiological integrity within and physiological autonomy from other organisms. Normal humans are the archetypal organism; other beings pass for organisms, even though they may be compromised on one or more of these three general features.

The Population Criterion

Populations follow from individual organisms, for they are collections of individuals. Even so, there are individuals that could be defined as populations in their own right. Colonies of primitive animals like sponges and anemone relatives have all the critical characteristics that we use to define a single organism. For the most part, populations contain one species, but this is by no means a requirement.

Populations vary in size from a few individuals chosen ad hoc for a given purpose, to millions of individuals spread over a large geographic area. Spatial discontinuity between populations is often a helpful criterion, but discontinuity can be very much a matter of degree. Even the convenient tangible bounds of an island can be a matter of degree, in that influx from surrounding populations combined with significant emigration could make the population as transient as a collection of human beings in a crosswalk. The physical bounds of the island might be perfectly clear, while at the same time they may correspond to no significant biological limit.

Often there are biological underpinnings to populations like a certain genetic homogeneity. This too is a matter of degree. Pollen comes on the wind to insert genetic variation in plant populations. Newcomers may readily invade animal groups. Species are special populations of large size with a degree of infertility in crosses between groups. For example, woodland primroses and their close relatives in the adjacent pastures regularly produce hybrids of intermediate form.[23] In animals, the circumpolar species of birds blend into one another, and interbreed with their neighbors.[24] The ends of the ring often cannot interbreed because the two end species are too divergent. Despite exceptions in populations of all sizes, populations can often be defined on grounds of relative genetic similarity within. Populations are collections of individuals that can be delimited by many criteria, depending on the question the ecologist chooses to ask. Organisms in populations are united with something of a shared recent history. Thus, a flock of different species may be a valid population. Recent interbreeding is only one of the histories that might apply.

The Community Criterion

Here we take a stance possibly at odds with convention. We do not insist on our point of view, and others are free to use their own definitions of communities. Research on the way ecologists conceive of communities shows us that plant ecologists do not, for the most part, work on communities using populations as the parts.[25] An analysis of the proportion of research papers on communities involving particular organisms forces us to the conclusion that the majority of ecologists conceive of communities as consisting of individuals rather than populations. Population studies disproportionately focus on animals and small plants. Trees can be studied as populations, but the fact that they are bigger than humans presses their individuality on us. That makes it hard to see populations in a forest. Animals can be readily seen to herd into populations, and the spatial limits of the populations of small plants can be observed easily by standing and looking at them from above. That is why animals and small plants are the favorites for population but not community work.

The explanatory principles of communities may involve competition, but not the clean, clear competition measured in population experiments. Competition in communities is set in a variable environmental context that does not allow population competition to come to exclusion.

Communities are the integration of the complex behavior of the biota in a given area so as to produce a cohesive and multifaceted whole that usually modifies the physical environment. The forest is an archetypal community with its individual trees bound together as members of the community by a tangle of processes. Trees standing side by side represent an instant in a continuing complex process involving not only competition but also interference, accommodation, and mutualism, among other factors. Our community conception only allows us to study competition, mutualism, and the many other population processes in aggregate. Communities are held together somewhat by microevolution. Species do not so much evolve into community membership as they represent multispecies units that arise as relatively stable entities coming from the ready-made species present. The species are like LEGO bits. The bits in isolation do not determine the community, but there are combinations of them that are stable. We see the stable configurations. The stability therein often may not be a product of evolution.

Because of the radically different space/time scales of animals and plants, coherence is generally only recognizable in either animal or plant communities, but not in combinations. Plants are good at plant taxonomy, and animals are good at animal taxonomy, but each is not so good at the other. This makes coherent response as multispecies assemblages difficult between kingdoms. There are plant and animal mixtures, such as biomes and coevolved species pairs or sets. In biomes, the insistence on species is relaxed, and is substituted by plant form. The form reflects physical manifestations to a shared climate. Animals groom biomes. Coevolved pairs of species are very robust. Communities come and go in each interglacial period, but species pairs, such as pollination partners, evolve together in, for instance, moth genera and orchid genera. Neither biomes nor coevolved plant–animal dependencies have a straightforward relationship to communities.

The Ecosystem Criterion

The functional ecosystem is the conception where biota are explicitly linked to the abiotic world of their surroundings. System boundaries include the physical environment. Ecosystems can be large or small. Size is not the critical characteristic; rather, the cycles and pathways of energy and matter in aggregate form the entire ecosystem. Robert Bosserman's thesis research was performed in the Okefenokee Swamp in Georgia.[26] His work was part of a big push on ecosystem research at the University of Georgia in Athens, led by Bernard Patten and organized by Edward Rykiel. Bosserman's research was part of ecosystem studies like others at the site. The difference was his bold reframing of ecosystems independent of size. Most everyone else was measuring big cycles of nutrients and spanning the large swamp, but Bosserman's work dealt with a natural microcosm. They were clumps of floating carnivorous plants called bladderworts, the size of a few handfuls. That plant performs photosynthesis and so provides the primary production. But being in a swamp, it lives in a low-mineral environment and turns to insects for its mineral supply. When a water flea touches a trigger hair, small bladders that are under negative pressure open and water with the flea in it rushes in. The flea is digested. Bosserman measured all the major functions in the clumps and discovered that all the major compartments of a typical ecosystem were present. Primary production, trophic consumption, and detritus compartments that fed production were all

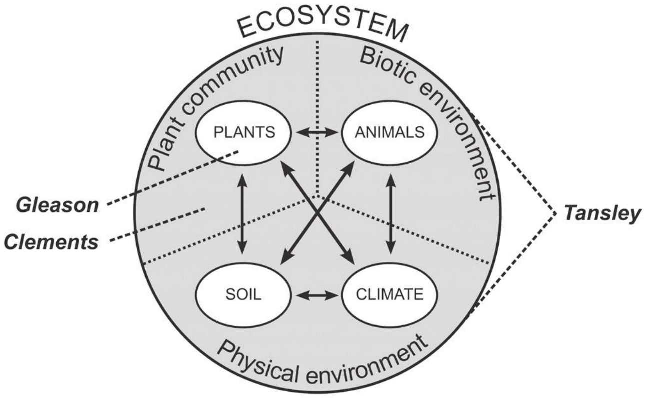

FIGURE 1.12. Gleason, Clements, and Tansley looked at vegetation in three different ways. Gleason saw the plant community as a collection of individuals filtered by environment. Clements saw the plant community as an integrated whole, set in a physical environment. Tansley saw the plants, their biotic environment, and their physical environment as all components inside the ecosystem. The same material system is seen from different perspectives.

there. Ecosystems are defined by flows and cycle, with the bladderwort operating as a natural microcosm, showing that ecosystems can be of almost any size.

The origin of the term "ecosystem" goes back eighty years to when Arthur Tansley recognized the need for an entity that blended biota with the physical environment.[27] That original explication persists so that conventional accounts of the ecosystem include a large box, the ecosystem, with four boxes inside labeled plants, animals, soil, and climate (figure 1.12). Unfortunately, such a characterization refers to the intellectual history of ecology in the 1930s more than to the powerful ideas that flowed from Tansley's brainchild. The ecosystem contains plants, animals, soil, and atmosphere, but those names are not helpful categories for seeing how ecosystem parts are put together. The functioning subunits in the ecosystem consist not of plants, animals, soil, and atmosphere, but of process fluxes between them. Plants do not naturally separate from soil when the compartment in question is "below ground carbon."

Tansley's term met a need of his day, for ecosystem approaches did precede the name itself. A full decade before the term ecosystem was coined, Transeau measured the energy budget of a cornfield.[28] He subtracted outputs from inputs to identify net gain. It would have been more productive had ecology taken that study as the archetype of an ecosystem. Transeau's implied ecosystem emphasized fluxes and pathways that are hard to address in an organism-centered conception that insists on preserving plants and animals as discrete entities.

The critical difference between an ecosystem and a community analysis is not the size of the study, but rather the difference in emphasis on the living material.

In community work, the vegetation may influence the soil and other parts of the physical environment by the very processes that are important in ecosystems. In communities, the soil is explicitly the environment of the plants. The biotic community is recognized as having integrity separate from the soil. The same vegetation on the same soil may be studied as a functional ecosystem or as a community, depending on how the scientist slices the ecological pie. Ecosystems are intractable if the biota is identified as one of the distinct slices, particularly if separate organisms are allowed to be discrete parts.

Take the case of the cycling of nutrients for repeated use. The leaf falls from the tree; worms eat the leaf; rainwater washes the nutrients into the soil directly from the leaf and from the feces of the worm; fungi absorb those nutrients and convey them to the root to which they are connected; the root dies, leaving a frozen core of nutrients; in spring, new roots grow down the old root hole, collecting the nutrients; and the rest of the plant passes them up to the leaves (figure 1.13). Harris, Kennerson, and Edwards inferred two periods of vigorous, below-ground growth in yellow poplar, once in the fall and then again in the spring.[29] In terms

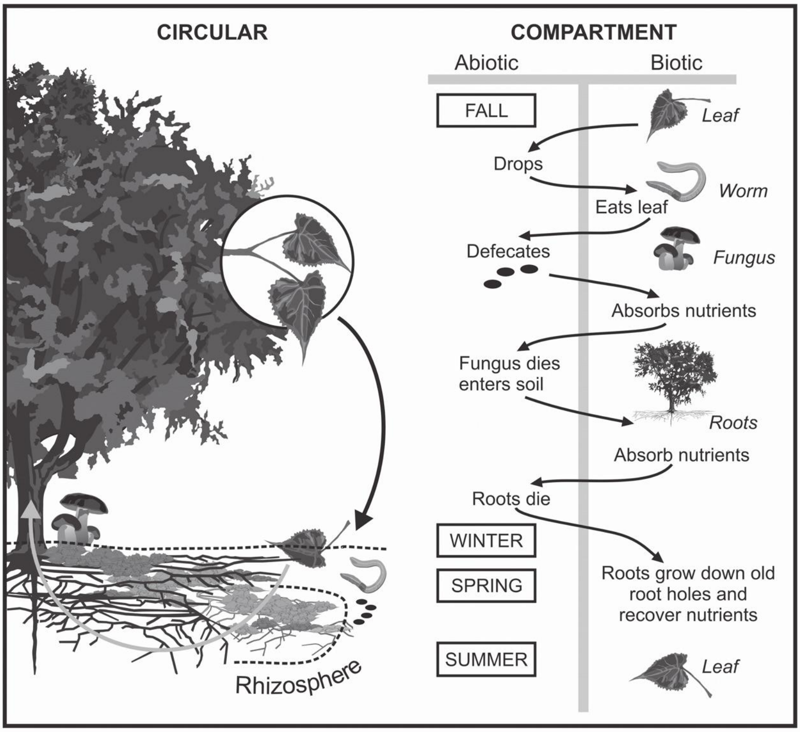

FIGURE 1.13. For a mineral nutrient to pass once around a nutrient cycle, it must pass in and out of the biotic component several times in one year. Accordingly, biota are not simple subsystems of process/functional ecosystems.

of a well-specified ecosystem, the earlier description is a simple, efficient nutrient pathway. In terms of a model that insists on emphasizing living material as different from nonliving material, the pathway is horrendously complicated. In just once around the cycle, the nutrients pass in or out of the biotic compartment at least four times.

The Landscape Criterion

Landscapes occupy a range of scales comparable to ecosystems and communities, but according to their own distinct criteria. Although landscapes are generally accessible from commonplace experience—and represent one of the earliest criteria for studying ecological systems—they have lain neglected for most of the twentieth century, ignored by almost all ecologists except those studying wildlife. The landscape ecology of the nineteenth century was supplanted by community conceptions.[30] Community ecology has its origins in an abstraction of landscapes; one where the pattern of the patchwork on the ground is replaced by abstract community types defined by species lists and proportions of species abundances. Pursuit of the community abstraction left the landscape conception untended. When a new conception was deemed necessary in the middle of the twentieth century, the ecosystem and not the landscape prevailed. In a sense, Victorian country vicars did all that could be done with landscapes before the computer age. In the last three decades, landscapes have again come to command the attention they deserve.

Spatial contiguity is the ordering principle for landscapes. As the dynamics of species replacement in community patches became a focus of study some forty years ago,[31] the scene was being set for a reawakening of interest in landscapes. More than in any other subdiscipline of ecology, landscape ecologists are buried in data. A casual glance across a vista reveals an unmanageable mass of detail. To achieve anything more than the obvious, the landscape ecologist must have cheap, fast computation. Recent years have seen the development of several measures of pattern that formalize and quantify what heretofore has been subjective, informal assessment. Remote sensing from satellites has allowed the landscape ecologist to move upscale. In many studies, the student of landscapes has only to look; that which made landscapes so obvious as to be trivial early in the last century now makes them natural objects for study in a modern, computer-assisted world. With tools to identify spatial structure, landscape ecologists now turn to fluxes of material in spatial contexts. Landscape ecology now investigates the consequences of elaborate spatial structure.

The Biome Criterion

By definition under conventional wisdom, biomes cover large areas. However, it is worth considering what the essence of a biome is, aside from size. Biomes are defined by the dominant vegetation physiognomy, something not strictly scale-defined in itself. A biome also should have a critical climate component such that the other characters are responses to some meteorological consideration. Often, an assemblage of animals plays a central role in giving the biome its particular structure. Examples here would be the spruce-moose biome or the grassland biomes with their respective grazers.[32]

As the name suggests, biomes are characterized principally by their biotic components, although soils and climate are important parts of the picture. Biomes, at first glance, are a hybrid of community and ecosystem with a strong landscape reference. The distinctive character of biomes is revealed when the concept is applied to situations scaled smaller than usual. Small systems that are simultaneously physiognomic, geographic, and process-oriented might prove very helpful. Using the scale-independent biome concept avoids the confusion that arises when we try to use one of the other criteria to describe such situations. Landscapes, communities, and ecosystems used separately or in tandem cannot do the biome concept justice. However, they are often pressed uncomfortably into service because we lack a term for small biomes. A frost pocket is a patch of treeless vegetation set in a forest. The absence of trees allows cold air to collect and kill any woody invaders. It is not adequately described as a community because the species are incidental to their life form. A frost pocket has all the regularly recognized biome properties except size: physiognomically recognizable, climate determined, disturbance created, and animal groomed.

The Biosphere

The biosphere is the one ecological system where the scale is simply defined. Its scope being the entire globe, it occupies a level that is unambiguously above all the other types of ecological systems discussed so far. Considering it a large landscape is possible but awkward. The biosphere is more often studied as a macrolevel ecosystem. For example, students of the biosphere ask questions about global carbon balance, a problem involving the same sort of fluxes considered by an ecosystem scientist. There is, however, a subtle but unequivocal shift in the relative importance of Tansley's four parts to the ecosystem when we make a shift upscale to the biosphere. The atmosphere is definitely an overriding influence in the biosphere, whereas in subcontinental ecosystems, soil interactions with biota through water are at least equal ecosystem players with the meteorological components. The biosphere is not easily conceived as a large biotic community because there is not interaction of species on a global scale to hold the global community together as a working unit.

We have looked at other criteria independent of scale, and solved the problem by identifying the means of investigation. Systems dominated by three-dimensional diffusion as the means of connection give plant physiological studies something in common with biospherics. Plant physiologists, such as Ian Woodward in

Sheffield,[33] often move into biospherics. He considers the density of plant stomata as a surrogate signal for the past global atmospherics.

Another feature of the global system is that it is close enough to a closed system to make workable the assumptions of closure that physicists are wont to make. So the ecology of closed systems has biospherical qualities. This is a critical issue for projects that put life in space. As Peter Van Voris and his colleagues were developing the microcosms they used to investigate ecosystem function,[34] some of the systems they probed were closed (the final ecosystem microcosms used were in fact quite open). In the closed systems, they found that those consisting of mostly liquid would boom and then crash, ending with a few species existing as resting stages on the wall of the microcosm. If they included a significant soil component in their closed systems, diversity in an active state would go on more or less indefinitely. They came to understand that the slower dynamics in the soil acted as an anchor to the rich functioning of the system.[35] Systems need a context for their functioning; otherwise, fast dynamics come to crash the system. The escape of the budworm in Holling's model shows that phenomenon (figure 1.11). The fast dynamics of insect outbreak booms and then crashes the system.

Biosphere II was an experiment on a closed system. Although there were problems with some of the science, it was a bold move to find out patterns that follow from a system being closed. In Allen's visit there, he saw many examples of small closed systems in sealed glass vials with persistent life in them, but as cysts on the glass walls. In the big Biosphere II, there were major biomes of varied sorts, to see if the functioning of the Earth turns on the diversity of forms that it contains. Once the system was closed, there was a steady depletion of oxygen. The reason was that the new concrete cured over time, taking in oxygen. As a result, there was some introduction of various materials, such as oxygen. One introduction was a copy of the first edition of this book. The biospherians were excited to get it. Allen gave a talk to the staff and the biospherians inside, and they all resonated to what it said. While there were problems with the original science in Biosphere II, the vision of it has been lost in a takeover by mechanistic science in recent decades, and few care anymore. Closure is a way to access the functioning of whole systems.

A FRAMEWORK FOR ECOLOGY

Biology is not just complicated chemistry and physics, it is a different discourse. Biology cannot deny the second law of thermodynamics; in fact, it depends on it.[36] We can see biology as a set of rules that work as a special case of physical happenings. Some limits in biology are physical, but many describe things that are possible in physics but not allowed in biology (figure 1.14). Nerves depend on potassium crossing a membrane by simple diffusion. But before the nerve fires,

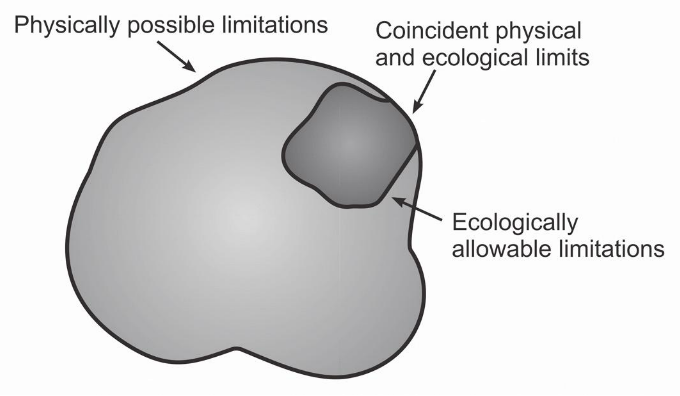

FIGURE 1.14. Only occasionally is what is observed in ecology directly the result of a physical limitation. More generally, what is ecologically allowable is a subset inside what is physically possible.

potassium is not allowed by the biology to cross the membrane and go down the density gradient; instead, it is actively moved in the other direction to become more concentrated on one side of the membrane. In all this, physics is not violated; it is just that some of what unconstrained physics will do is not allowed.

Physics does not have precedence or privilege. In fact, Robert Rosen would say that the reverse applies. In his classic work in biology, *What Is Life,* Erwin Schrödinger pointed out that there was not a physics that applied to biology. And Rosen submits that such still pertains. The reason is that physics, as it is practiced, needs to make a set of assumptions. Now, those assumptions in physics—like all assumptions everywhere—are not true. Systems are not in fact closed, but in physical systems you can get away with calling them so. The point is not to find which assumptions are right, but what lies you can get away with. You simply cannot get away with assumptions of closure in biology. And there is on its face no life close to equilibrium. At equilibrium, life dies. Rosen says we should put open, far-from-equilibrium biology, such as organisms, in physics through analogy.[37] Then we will have physics for biology.

As scientists, we take a position that is distinctly data driven. When we recognize a type of thing, we are scanning our past experience for things that are like what we see. So type is subjectively chosen by us, but in the light of experience. Do not imagine then that type is only internal to ourselves, it involves things we have already experienced. Also, do not imagine that the scaling of things we see is imposed on us by nature, with no values on our part. Consider a pond skater and a big dung beetle. Put both on the surface of water, and one skates and the other sinks. It is a matter of scale; the beetle is too big to skate. Now put a metal bolt on the water; it too will sink. But somehow that is not a matter of scale; it is simply a different situation. We as biologists would dismiss the bolt as something else, not something heavier. But let us ask a physicist, who we would imagine saying, “Oh, it is a scale issue. The bolt is heavier than the skater so it breaks the surface tension, just like the beetle.” Biologists might say that the beetle and the skater are both insects, and so are similar enough to have one scale applied to them. The bolt is not similar enough for our biological sensitivities. Thus, scale or not scale depends on the observer, and so scale has some of its origins inside us and our decisions.

The subjectivity of naming things ecological is crucial to ecological understanding. Calling something an ecosystem involves taking a point of view on a given tract of land that emphasizes something different than if one called that same material system a community. The framework we erect subjectively defines entities of an ecological sort and embeds them in a physical setting where scaling comes to the fore.

Ecological systems are complex and require careful analysis if the student of ecology is to avoid being lost in the tiered labyrinth of the material. The most important general point covered thus far is the recognition of the role of the observer in the system. Ignoring human subjectivity will not make it go away. Since one makes arbitrary decisions anyway, ignoring them abdicates responsibility needlessly. All decisions come at the price of not having made some other decision. By acknowledging subjectivity, one can make it reasoned instead of capricious. A real danger in suppressing the ecologist in ecology is to be bound by unnecessarily costly decisions. These could be exchanged for a more cost-efficient intellectual device if only the subjectivity of the enterprise were acknowledged.

One of the most important benefits of consciousness of the observer is knowing the effects of grain and extent on observation. The formal use of grain and extent gives opportunities to avoid old pitfalls and opens up possibilities for valid comparison that have heretofore escaped our attention. Also of great importance is being cognizant of patterns of constraint. Constraint gives a general model for couching ecological problems in terms that generate predictions. At this point, we have put most of the crucial tools in the box and are ready to go to work. It is time to erect a general frame inside which we hope ecology will become a more predictive science. It will certainly be more unified than the fractious thing that is generally practiced at the time of this writing.

Given the extended treatment of definitions of ecological criteria and the tools for linking them put forward in this chapter, the framework we suggest can now be adequately stated in a short space. The temptation we resist is to stack types of ecological systems according to an approximation of their size. Yielding to the desire for tidiness in the conventional ecological hierarchy is costly.

Up to this point, we have detailed our paradigm describing the difference between scale-based levels of observation from levels of organization. Our representation for this paradigm is a cone-shaped diagram containing columns to represent ecological criteria at different scales (figure 1.15*A*). Instead of stacking criteria one upon another in the conventional order, the diagram separates scaled ordering from an ordering of levels of organization. As a result, we can hold the scale constant and look at how the criterion under consideration is but one conception of a material system at a given temporal and spatial extent, for example, considering a forest as a patch of landscape, as a community, or as an ecosystem. We can put each type of ecological system at every scale. Organism, population, community, ecosystem, landscape, and biome systems will sit side by side at every scalar level. Not only can we compare the ecosystem and community conceptions

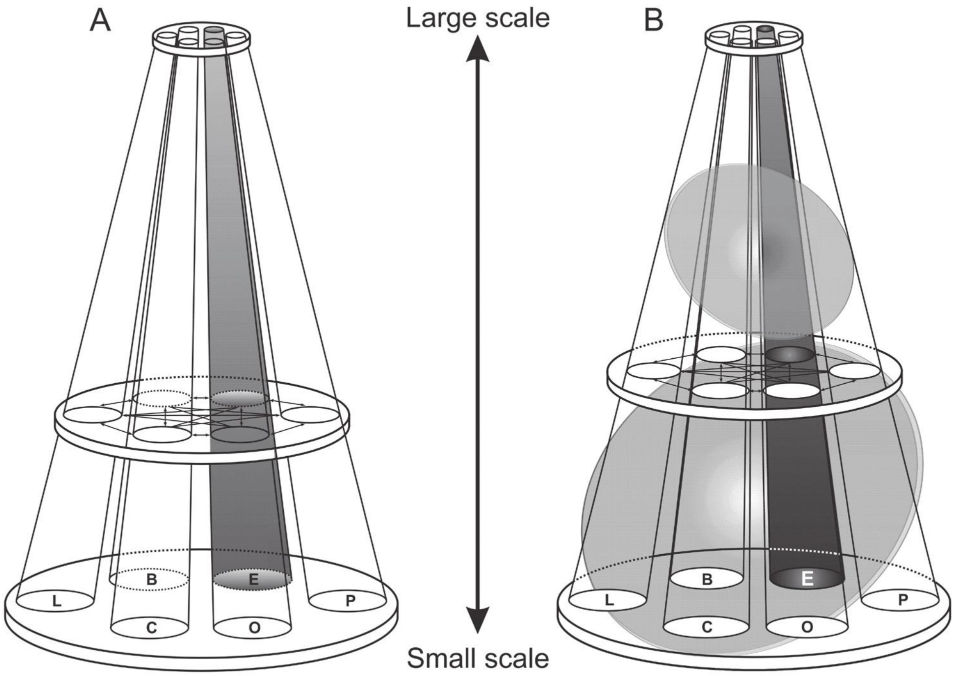

FIGURE 1.15. *A*. The cone diagram metaphor for ecological criteria and ecological scale. The wide base indicates a large number of small entities; the narrow top indicates a small number of large entities. The cross section across the entire cone represents one middle-level scale. Although there is only one here, any number of cross sections could be inserted, each at its own scaled level. Each letter indicates a different criterion: O = organism; P = population; C = community; E = ecosystem; L = landscape; B = biome. The six lettered disks correspond to a comparison of scale-independent criteria that intersect at a given grain and extent. Individually, the columns represent a criterion for looking at the material system, for example, the abstract notion of community. The disk labeled C is an actual community with a particular spatiotemporal size assigned to it. In the C column, larger-scale contextual communities occur above that community, while smaller-scaled community subsystems would occur below. *B*. The large diagonal disks show an ecosystem, as an example, of a given size working as the context for a smaller community, while itself being held in the context of a larger community. Many other diagonal cross sections are possible, such as one that sees a cow's rumen as an ecosystem, obviously scaled below the cow itself. The organism is the context of its rumen ecosystem.

across a landscape of a given area, but we can do the same at larger- and smaller-scaled landscapes.

Our scheme is not limited to making comparisons horizontally across a given spatiotemporally defined level. We can move up and down the scale on the cone diagram and make comparisons between differently scaled entities of a single type of system. We could see how a given ecosystem contains smaller ecosystems, while being itself part of a larger ecosystem. As we move to levels below, we may find the mechanistic explanations of the behavior of the level in question. The levels above define the role of entities at the level in question in the functioning of larger systems. The upper levels define role, purpose, and boundary conditions. Predictions come most readily when the system to be predicted is up against a constraint imposed by the layer above.

Probably the most interesting—and certainly the most neglected—questions will involve slicing the cone diagram diagonally (figure 1.15*B*). Here, we change the type of system while also changing levels. In a sense, that is what the conventional hierarchy does, but we are not limited to the conventional order. Any type can be the context for any other type. Diagonal slices will allow us to see how community patterns at a given level are influenced by nutrient status maintained by a subsystem defined in ecosystem terms. With our ecological cone diagram, we can ask how low-level ecosystem function is constrained by population considerations of a dominant tree population. We no longer need to see a difficulty when animal communities carry nutrients around an ecosystem that sits across the boundary between two biomes. The biome boundary is clear by its own criteria, but it is leaky with respect to nutrients involving animals and ecosystems. We do not recommend a compromise that forms a general-purpose system designed by a committee composed of one population biologist, a community ecologist, an ecosystem scientist, and a biogeographer. Rather, we suggest a formal change in the type of system description every time a new explicit question or explanation demands it. We use the above scheme to organize the material in the chapters that follow. We see it as both flexible and encouraging consistency.

The cone diagram of figure 1.15 is our version of the Beaufort scale. We use it to fix order in rich ecological situations that are at least as complicated the seas that Beaufort assessed. Beaufort only separated situations on wind speed and its consequences. Our use of distinction between type of situation and its spatiotemporal scaling gives a richer device, one more suited to the many nuances of ecology.