

The American Naturalist
---An empiricist's guide to using ecological theory
--Manuscript Draft--

| | |
|--|--|
| Manuscript Number: | 60651R2 |
| Full Title: | An empiricist's guide to using ecological theory |
| Short Title: | Empiricist's guide to theory |
| Article Type: | Synthesis |
| Additional Information: | |
| Question | Response |
| Data Accessibility during Peer Review. Data should be accessible during peer review. Files can be uploaded with your paper file or deposited for double-blind peer review in Dryad. Please provide the temporary Dryad URL (NOT the DOI) or other location of your data (or indicate there is no data). The information in this field will be provided to reviewers. | The code used to produce Figure 2 has also been archived with Zenodo (https://doi.org/10.5281/zenodo.5154139) |

Title: An empiricist's guide to using ecological theory

Authors: Tess N. Grainger,^{1,2†} Athmanathan Senthilnathan,^{3*} Po-Ju Ke,^{1,4*}

Matthew A. Barbour,⁵ Natalie T. Jones,⁶ John P. DeLong,⁷ Sarah P. Otto,² Mary I. O'Connor,²

Kyle E. Coblenz,⁷ Nikunj Goel,⁸ Jawad Sakarchi,² Megan C. Szojka,^{2,9} Jonathan M. Levine,¹

and Rachel M. Germain^{2†}

*These authors contributed equally

†Corresponding authors: tessgrainger@gmail.com and rgermain@zoology.ubc.ca

Author affiliations: 1. Princeton University, 2. University of British Columbia, 3. University of Tennessee, 4. National Taiwan University, 5. University of Zurich, 6. University of Queensland, 7. University of Nebraska-Lincoln, 8. University of Texas at Austin, 9. University of Wyoming.

Running head: Empiricist's guide to theory

Keywords: hypothesis testing, experiments, experimental design, methodology, theory, mathematical models

MS type: Synthesis

Abstract

A scientific understanding of the biological world arises when ideas about how nature works are formalized, tested, refined, and then tested again. Although the benefits of feedback between theoretical and empirical research is widely acknowledged by ecologists, this link is still not as strong as it could be in ecological research. This is in part because theory, particularly when expressed mathematically, can feel inaccessible to empiricists who may have little formal training in advanced math. To address this persistent barrier, we provide a general and accessible guide that covers the basic, step-by-step process of how to approach, understand, and use ecological theory in empirical work. We first give an overview of how and why mathematical theory is created, then outline four specific ways to use both mathematical and verbal theory to motivate empirical work, and finally present a practical toolkit for reading and understanding the mathematical aspects of ecological theory. We hope that empowering empiricists to embrace theory in their work will help move the field closer to a full integration of theoretical and empirical research.

Introduction

“The great book of nature is written in the language of mathematics.”

– Galileo Galilei c. 1600

Scientific inquiry should operate as a feedback loop in which theory that describes the natural world is developed, tested empirically through carefully articulated hypotheses, modified to better represent reality, and then tested again. When this feedback loop works, theory provides a framework to guide inquiry, experimental design, and the interpretation of observed patterns, supplies mathematical tools to harness information from collected data, and connects individual experiments to general ideas about how nature operates. In turn, empirical research can be used to support, refute, or revise theoretical predictions, indicate which theoretical assumptions are consistent with the natural world, and point theoreticians to overlooked processes that can be integrated into models.

Feedback between theory and experimentation has always been essential for the progress of ecology (Gause 1935; Huffaker 1958; Simberloff and Wilson 1969; Tilman 1977), and as collaborations across disciplines become more common, theoretical and empirical research have the potential to be more tightly linked than ever before (Laubmeier et al. 2020; Rossberg et al. 2019; Servedio 2020). Moreover, in confronting global challenges such as climate change, biodiversity loss, and emerging diseases, humanity can benefit enormously from the ecological understanding that arises from a fully functional scientific process (Ferrier et al. 2016; Marquet et al. 2014).

Unfortunately, the links between theoretical and empirical research in ecology are not as strong as they could be (Caswell 1988; Fawcett and Higginson 2012; Haller 2014; Kareiva 1989; Rossberg et al. 2019; Scheiner 2013). Up to 45% of empirical ecology papers make no mention of any theory whatsoever (Scheiner 2013), and fewer than 10% of ecologists and evolutionary biologists agree with the statement that “theoretical findings drive empirical work” in their fields (Haller 2014). This disconnect has been attributed to several underlying challenges, including a lack of emphasis on theoretical training in ecology (Rossberg et al. 2019), a lack of motivation on the part of some theoreticians to engage with the language of empiricists (Grimm 1994) or with the elements of nature that empiricists focus on (Krebs 1988), a general lack of mutual appreciation between empiricists and theoreticians (Haller 2014), and persistent communication barriers between these two groups (Servedio 2020). A major cause of the communication barriers between theoretical and empirical ecologists is that in ecology, as in other scientific fields, theory is often expressed in the language of math. Ecologists come to the field through a variety of educational and experiential pathways, are often drawn to ecological research by an interest in natural history or hands-on field and lab research, and may not be formally trained in advanced math. Likewise, theoretical research may not be written for a general audience and can contain unexplained assumptions, terminology or notation that hinders comprehension (Dee-Lucas and Larkin 1991; Grimm 1994). As a result, the mathematical aspects of theory can feel inaccessible to empiricists (Fawcett and Higginson 2012; Haller 2014). Indeed, the more equations an ecology and evolutionary biology paper contains, the fewer citations it receives (Fawcett and Higginson 2012). Of the empirical papers that do cite mathematical theory, only about half do so both correctly and specifically (Servedio 2020). This barrier presents a major challenge to the full integration of theoretical and empirical work in ecology.

Previous papers and books have discussed the causes and extent of the disconnect between theoretical and empirical research in ecology (Fawcett and Higginson 2012; Haller 2014; Kareiva 1989; Łomnicki 1988; Rossberg et al. 2019; Scheiner 2013). The value of theory in ecology and evolutionary biology has been emphasized, and many have advocated for a better integration of theory into empirical work (Caswell 1988; Marquet et al. 2014; Pickett et al. 2010; Servedio 2020; Servedio et al. 2014). Detailed, step-by-step instructions for building biological models and fitting them to data have been provided (Bodner et al. 2020; Bolker 2008; Otto and Day 2007; Pielou 1969). Recent work has helped to demystify the process of theory creation (Otto and Rosales 2020), has advocated for the training of more theoretical ecologists (Rossberg et al. 2019), and has provided an overview of some of the challenges associated with communicating theory to a broad audience (Shoemaker et al. 2021). However, this still leaves many empiricists who are interested in motivating their research with theory, but have little prior experience with math and models, struggling to take the first step.

We seek to fill this gap by providing an accessible guide that covers the basic, step-by-step process of how to approach, understand and use ecological theory in empirical work. We achieve this by (1) briefly reviewing how and why mathematical theory is created, (2) outlining four specific ways that mathematical and verbal theory can be used to motivate empirical work, and (3) providing a toolkit for reading and understanding the mathematical aspects of ecological theory. We are ecologists and evolutionary biologists spanning the full spectrum from empiricists to theoreticians, and this guide is the outcome of the empiricists amongst us feeling intimidated by the mathematical aspects of theory and reaching across the divide to our theoretical colleagues to build the toolkit needed to understand and use theory. While we recognize that the onus should not fall entirely on empiricists to bridge the theory-empirical

divide, and that fostering collaborations between theoreticians and empiricists and encouraging theoreticians to present their theory in a clear and accessible way are also critical for achieving this goal, in this paper we focus on the steps that empiricists can take. Our aim is not to teach empiricists how to create theory or to express it as math, but rather to empower empirically-minded ecologists to read and understand mathematical theory and to integrate theory into their work at every step in the research process.

A primer on theory

We begin by introducing the general concept of theory: what it is, what it aims to do, and how it is created. While we focus on ecological theory, many of the ideas discussed here will also apply to evolutionary theory and to other scientific fields. We focus primarily on theory that is expressed mathematically because these elements of theory often present the steepest accessibility barrier to empiricists, while recognizing that theory and math are not inexorably linked. Indeed, many excellent theories do not involve math (e.g., *The Theory of Evolution by Natural Selection* (Darwin 1859)), and many uses of math in ecology are not theory (e.g., practical applications of statistics). Finally, we restrict our discussion to model development that does not use data, but discuss statistical model fitting as an empirical approach (see ‘Use the mathematical equations’).

What is theory?

We define **ecological theory** broadly, as an explanation of an ecological phenomenon. These explanations take the form of narratives that explain how an ecological process works or why an ecological pattern is observed, and that become scientifically useful when expressed in a logical structure (Pickett et al. 2010; Rossberg et al. 2019). The transformation of an idea in narrative form into logical testable theory often, though not always, involves the use of mathematical

models (Otto and Rosales 2020). A **mathematical model** is an equation or a set of equations that describes how different aspects of a system relate to one another (Otto and Day 2007). The term ‘**model**’ refers to the fact that mathematical models are idealized and simplified versions of reality, just as architectural models represent key features of complex structures and model organisms represent a group of organisms that share common attributes (Kokko 2007). The translation of ideas in a theoretician’s mind into math on a page serves several purposes: math provides a clearer and more objective expression of relationships, it brings to light assumptions and logical errors that may be obscured in verbal hypotheses, and it places ideas and hypotheses into a concrete and concise form (Grimm 1994; Kokko 2007; Marquet et al. 2014; Otto and Rosales 2020).

How is mathematical theory created?

The process of expressing theory in mathematical terms can seem mysterious to those who have never attempted it, which can make theory seem unapproachable. Here we pull back the curtain on theory creation by describing the typical steps that a theoretician takes when creating a mathematical model (Bolker 2008; Otto and Day 2007), and in doing so, highlight the many similarities between conducting theoretical and empirical research.

Step 1: The motivation. The first step in theory creation often parallels that of empirical research: the theoretician thinks of a biological question that she is interested in or a biological process that she wants to understand. The question may be motivated by an unsolved problem; for example, how temperature affects consumer-resource dynamics (DeLong and Lyon 2020; O'Connor et al. 2011). Or it could be motivated by an empirical observation; for example, evidence of interspecific differences in plant species’ susceptibility to pathogens and an interest in the implications for plant coexistence (Ke and Wan 2020; Stump and Comita 2018). Or it may

be motivated by a practical problem; for example, the need to understand how population structure influences COVID-19 dynamics (Britton et al. 2020).

In other cases, the creation of new theory is motivated by the desire to extend an existing model rather than to develop a new model from scratch. Indeed, most models are variants of previous models, extended to include different biological features, to tailor the model to a new system, or to address a slightly different question. One underappreciated initial goal of some theory is to provide a **proof-of-concept** that verifies or refutes ideas that may have previously only been expressed verbally (Servedio 2020). This use of theory demonstrates that something can happen under certain conditions, but not necessarily that it is likely to occur under all, or even most, conditions (Otto and Rosales 2020; Servedio 2020; Servedio et al. 2014). These models can serve as hypothesis tests themselves and may not be meant to be tested empirically (Caswell 1988; Servedio et al. 2014). More broadly, it is important to note that creating theory and translating it into math is a means of scientific discovery unto itself, and while in this paper we focus on the ways in which empiricists can use theory, theory need not be tested empirically in order to be useful. Theory can have a major influence on a field without any empirical testing, for example theory demonstrating that transient dynamics can be prevalent and long-lasting (Hastings 2001). Recognizing that not all theory was necessarily meant to be tested empirically can help empiricists focus on the types of theory that are most relevant to their scientific goals.

Step 2: The method. Next, the theoretician looks at the approaches that theory addressing similar questions has used and decides whether these approaches are relevant, justified, and suitable. The approach a theoretician takes is likely to be guided and constrained by her background; for example, a theoretician who is most familiar with **dynamical models** (those that track how a system changes over time) may be drawn to research questions that can be answered

by that type of model, just as an empiricist who works with plankton mesocosms may focus on questions that are well-suited to that approach.

A feature of this stage of theory development that may not be obvious to empiricists is the variation in the degree to which a new model builds on, tweaks, and combines existing models, versus conceiving of an entirely new mathematical approach. For example, parameterizing predator-prey population dynamic models with expressions for temperature-dependent biological rates from metabolic theory led to new insights about how temperature affects the abundance of interacting consumers and resources (O'Connor et al. 2011). Likewise, many COVID-19 models are variants of the classic “SIR” epidemiological model documenting transitions among Susceptible, Infectious, and Resistant hosts. In contrast, mathematical approaches drawn from physics were adopted by ecologists to understand the complex dynamics of animal grouping behavior (Okubo 1986), and math from information theory have yielded new perspectives on ecological problems such as energy flow in food webs (Rutledge et al. 1976) and biodiversity (Jost 2006). For empiricists interested in creating models, the recognition that theory is often created by combining existing models in new ways can help reduce the perceived barriers to participation.

Step 3: The equations. In this step, the theoretician first decides which parameters and variables are necessary to describe the biological process of interest. A **parameter** is a quantity that has a value that does not vary within a given equation (e.g., resource supply rate), while a **variable** is a quantity whose value can vary within the equation, for example through time (e.g., population size). In the simple model describing a linear relationship between x and y ($y = mx + b$), x and y are variables (there are multiple values of x and y), and m (the slope) and b (the y intercept) are parameters (there is only one slope and one y intercept for this equation). Note that while

parameters do not vary within an equation, the biological processes that they represent do vary in nature (e.g., across space, time, organism identity) and changing the value of parameters and exploring how that affects outcomes of interest is a common exercise in theory development (see “numerical approaches” below). A helpful exercise when deciding how multiple related equations should be written out is to draw “stock and flow” diagrams that show the relationship between the different variables in each equation (Otto and Day 2007).

In determining the equations, the theoretician must make difficult choices between how broadly the model can be applied across systems (generality), how closely the model mimics real world processes (realism), and whether the model produces specific quantitative predictions or general qualitative predictions (precision) (Bodner et al. 2020; Levins 1966; Servedio et al. 2014). For example, models can be written with more terms or relationships to capture more biological detail, but additional terms may render the model less general. Note that the aim of mathematical models is to increase our understanding of some phenomenon, not to describe nature perfectly, and so a model making unrealistic assumptions or simplifying a process is not necessarily a failing of the theory or a reason to discount its ability to guide empirical research (Grimm 1994; Kokko 2007; Levins 1966; Phillips 2015; Pielou 1969; Servedio 2020). Just as empirical studies vary widely in their realism, from field-based observational studies to highly manipulated lab experiments, so too do ecological models. And just as each type of empirical study has value, so too does each type of theory.

Step 4: The analysis. The theoretician then uses mathematical procedures to understand the behavior of her equations within the range of parameters that she is interested in exploring. This can include solving for equilibria, investigating the stability of the equilibria, or studying how a response of interest (e.g., equilibrium density) varies across a range of parameters. These are

common ways of determining the “outcome” of the processes being modeled. Models can be analyzed analytically or numerically. **Analytical approaches** use mathematical techniques to solve the equation (i.e., find the equation that has the output of interest on the left-hand side) (Table S1). This approach requires mathematical labor but produces more general conclusions (Otto and Day 2007). **Numerical approaches** (e.g., “simulations”) involve setting model parameters to specific values and observing the outcome graphically or as numeric values (Table S1). Here, conclusions are limited to the specific range of parameter values being explored, but because no analytical solutions are needed, theoreticians may turn to this approach when a model is too complex to be approached analytically. Numerical approaches are somewhat akin to empirical experiments in that the theoretician varies a specific process (e.g., resource supply rate) to determine how that affects an ecological outcome (e.g., consumer population density). While analyzing a model’s behavior cannot match the biological realism of an empirical experiment, it does allow researchers to hold potentially confounding factors constant and to explore a wide range of parameter space, whereas empirical experiments need to randomize factors that cannot be controlled (e.g., individuals, sites) and are often logistically constrained to smaller ranges or discrete categories (e.g., “low” vs “high”). Likewise, the ability to depart from the realities and constraints of the natural world allows theoreticians to explore questions that would be difficult or impossible to address empirically, such as the consequences of very long time scales, complex interactions or extreme trait values.

Step 5: The end. Finally, the theoretician refines her equation(s) and writes up and publishes the research questions, model, analyses, and conclusions. To refine the equations, the theoretician will often check whether the results make sense, both logically and in light of previous empirical or theoretical work, and revise the model and analyses accordingly. In writing up her findings,

the theoretician will ideally follow practices that help empiricists understand and use her work, such as describing how the model follows or departs from previous theoretical research, interpreting the results verbally, including a table of all model parameters and their biological meanings, and describing experiments that could be used to test the predictions. A full integration of theoretical and empirical research in ecology will depend on sustained efforts by both empiricists and theoreticians to bridge the divide, and these efforts will be most fruitful when theoreticians are motivated to write up their findings with empiricists in mind and to highlight assumptions and results that can be measured and tested in the lab or field.

How to use theory in empirical research

In this section we outline four ways that theory can be integrated into empirical research. These classifications are not strict categories, and a single theory may fall into several categories. This organization simply represents one way to think about theory that can guide empiricists as they move through different stages of their research. All four of these approaches could be used in a single study, as they are applied at different stages of project design (Fig. 1).

An important philosophical question to consider when integrating theory into empirical work is what “testing a theory” actually means. To us, an empiricist tests theory by using empirical data to evaluate whether the outcomes predicted by the model or the assumptions it makes occur in a natural system, not whether the math is correct (Servedio et al. 2014). Two of the approaches covered below (‘Test the predictions’ and ‘Test the assumptions’) involve “testing” theory. The other two approaches (‘Adopt the framework’ and ‘Use the math’) do not actually test the theory, but rather accept that it is applicable enough to be useful in guiding empirical research.

Approach 1. Adopt the framework

Some ecological theories act as general frameworks for understanding ecological dynamics (in the broadest sense) or for understanding a major class of ecological processes (Pickett et al. 2010). While these theories may also provide specific mathematical equations or predictions, the larger contribution is often a new way of thinking about a problem or a unification of related ideas. Examples of this type of theory include The Theory of Evolution by Natural Selection (Darwin 1859), Consumer-Resource Theory (Lotka 1926; Volterra 1927), Grime's Triangle of Plant Strategies (Grime 1977), Life History Theory (Pianka 1970), Modern Coexistence Theory (Chesson 2000), The Neutral Theory of Biodiversity (Hubbell 2001), The Metabolic Theory of Ecology (Brown et al. 2004), and The Maximum Entropy Theory of Ecology (Harte and Newman 2014). Some of these theoretical frameworks are purely verbal, some are heavily mathematical, and some involve a mixture of both verbal and mathematical reasoning.

Theoretical frameworks create a scaffold within which ecological patterns and the processes that generate them can be understood, and they can provide the context in which to integrate evidence from experiments, observational studies, and mathematical modelling (Fig. 1). In some cases, a new theoretical framework sparks the interest of empiricists and prompts a rapid rise in empirical work focused on a certain topic.

How to do it

Theoretical frameworks can enhance empirical research in several ways: they can reorient how we understand and study a biological process (e.g., The Theory of Evolution by Natural Selection), they can unite disparate lines of evidence into a unified framework (e.g., effect of temperature across organizational scales in Metabolic Theory), they can help focus empirical research on a specific process or relationship (e.g., dispersal in Metapopulation

Theory), or they can provide a null model for how patterns emerge in nature (e.g., The Neutral Theory of Biodiversity). Although theoretical frameworks can be extremely useful, it is also important to keep in mind that all theoretical frameworks have shortcomings, that many will eventually outlive their utility, and that the lack of diversity amongst the authors recognized for generating the theoretical frameworks in ecology has undoubtedly limited the range of perspectives represented.

A useful first step when motivating empirical research with a theoretical framework is to become familiar with the framework and with existing empirical work that has adopted it. This can help clarify which aspects of the framework have been well-explored and where gaps in our collective understanding remain. It can also help to reveal common experimental designs that address the questions or capture the processes described in the framework (Broekman et al. 2019; Grainger and Gilbert 2016). This step can also help reveal situations in which new empirical methods are needed to apply an emerging framework, which can be an exciting opportunity for empiricists (e.g., methods to quantify selection on phenotypic traits on ecological timescales introduced in Lande and Arnold (1983); methods to account for phylogenetic relatedness in comparative studies introduced in Felsenstein (1985); methods to apply Modern Coexistence Theory introduced in Godoy and Levine (2014)).

Approach 2. Test the predictions

Theoretical work often generates specific predictions, and testing these predictions is a common way that empiricists use theory (Fig. 1). The central aim of testing predictions from theory is to determine whether a pattern that a theoretical model predicts matches a pattern that manifests in a natural or experimental system, ideally with the end goal of using the theory to deepen our understanding of the processes that create those patterns. Examples of recent

empirical tests of classic theoretical predictions include the predicted increase in species richness with increasing island area and connectivity from the Theory of Island Biogeography (proposed by MacArthur and Wilson (1967); e.g., tested by Prugh et al. (2008)), the predicted negative density dependence of seed recruitment from the Janzen-Connell Hypothesis (proposed by Janzen (1970) and Connell (1971); e.g., tested by Harms et al. (2000)) and the predicted strategy of delayed germination in variable environments (proposed by Cohen (1966); e.g., tested by Venable (2007)).

How to do it

The first steps in testing theoretical predictions are to consider what type of model the theory uses (Box 1), to explore the model and understand (well enough) what the math means biologically (see ‘A toolkit for understanding equations’), and to determine what assumptions the theory makes (Box 2). Next, the empiricist can consider how the theory can inform experimental design. The theoretical paper will ideally lay out its predictions clearly, in words and in figures (see Box 3); for example, the prediction that local diversity is maximized at intermediate levels of dispersal between habitat patches (Mouquet and Loreau 2003). The empiricist can then determine what needs to be manipulated and what needs to be measured: here, an experiment in which organisms’ dispersal rates between local patches are manipulated and diversity is measured (e.g., Parain et al. (2019)). Theory can also help inform the more detailed elements of experimental design, for example: should all organisms have the same dispersal rates? How many dispersal levels are needed to detect the predicted relationship? Is repeated sampling over time necessary? For each question, the empiricist will have to consider trade-offs between matching the experimental design to the theoretical assumptions, matching the design to the

attributes of a natural system, and designing an experiment that is feasible to implement (Box 2) (Grainger and Gilbert 2016; Laubmeier et al. 2018; Uszko et al. 2020).

At this stage, a careful consideration of what evidence is needed to strongly support or refute a theoretical prediction can be valuable. Powerful empirical tests of theoretical predictions document a pattern (or ideally multiple patterns) that are very unlikely to occur by chance or alternative mechanisms, that support a single hypothesis, and that (ideally) reject one or multiple alternative hypotheses (Loehle 1987; Platt 1964). They quantitatively measure responses to the manipulation of the independent variable, they have sufficient replication to detect the theoretically predicted responses, and they use experimental conditions that are consistent with the theory's assumptions. Doing so makes it possible to draw conclusions about the validity or generality of the theory based on whether or not results are consistent with the theory's prediction.

However, a challenge with hypothesis testing in ecology is that the context and system-specific nature of ecological dynamics mean that results that fail to reject a null hypothesis are not necessarily a strong refutation of a theory (Hardin 1960; Loehle 1987). As such, it is worthwhile to carefully consider what conclusions can be drawn if experimental results do not support the theoretical prediction, and how the experiment can be designed so that any outcome will provide interesting insight into the process of interest. One way to conduct more nuanced hypothesis testing is to ask: "*when* does x affect y, as predicted by theory?" rather than "*does* x affect y?" The answer is more likely to be "sometimes" than "always" or "never." For example, research suggests that dispersal is a stronger driver of diversity in the absence of predators than when predators are present (Chase et al. 2010; Kneitel and Miller 2003). There is a good reason why many conclusions in ecology include the word "can": it takes a critical mass of empirical

evidence from many systems in order for the field to accept that a theoretical prediction is broadly and generally supported. However, while contributing empirical support for a theoretical hypothesis is a worthwhile achievement, we also strongly encourage the dissemination of **negative results** (those that detect no effect), which make important but undervalued contributions to scientific progress and can highlight areas where existing theory needs modification (e.g., Fig. 1) (Knight 2003; Loehle 1987).

Approach 3. Use the mathematical equations

Sometimes a valuable part of a theory to an empiricist is a specific equation that can be used to obtain a quantitative estimate of a biological process that is difficult or impossible to measure directly (Fig. 1). For example, measuring the strength of competitive interactions or the rate of disease transmission is difficult, but established models can be used to estimate these from more easily measured quantities such as birth and death rates. Below we discuss two specific ways that an empiricist can use equations from theory to gain new information: model fitting (in which parameters are estimated), and the direct use of an equation (in which the response variable is calculated). A third and very straightforward way that an empiricist can “use” mathematical equations, particularly those that are tailored to a given biological system, is to help determine which biological processes are most relevant to a particular outcome and should therefore be measured or manipulated in experiments. For example, models of flour beetle dynamics identify several types of cannibalism as important to population growth (Dennis et al. 1995), while annual plant models highlight germination rate and seed bank viability as key mediators of population growth rates (Levine and Rees 2004). This guidance can be particularly helpful for empiricists starting work in a new system.

How to do it

The first step in using equations is to figure out what information is desired and which candidate models can provide that information. At this stage, it is important to check that the assumptions of candidate models are satisfied by the biological system and experimental design being used (Box 2), because once a model is used to estimate a biological process, it becomes implied that the system meets the assumptions of that model, and the quality of any subsequent results will depend on this being (mostly) true (Clark et al. 1998). When selecting candidate models, established models can be used or existing models can be modified to reflect the dynamics and conditions of the study system. The modification of existing models to fit the study system or the dynamics of interest can blur the line between using and creating theory, which is often a sign of effective integration of empirical and theoretical research (e.g., Duffy et al. (2005); Moeller et al. (2016)).

The first way to use equations in empirical research is “**model fitting**” (a.k.a “fitting models to data”). To do this, the empiricist identifies mathematical models that describe a process of interest, collects measurements of the response and predictor variables in these models, and then uses statistical techniques to estimate the values of **free parameters** (those whose values are unknown) that best match the relationship observed in the data (Bolker (2008) and Table S1). Each of the free parameters represents a biological process or property of interest. An example of this approach is using a functional response model to estimate two parameters, the space clearance rate (a.k.a attack rate) and handling time in a predator-prey interaction, by inputting data from an experiment that quantifies the number of prey consumed (response variable, y) when different numbers of prey are available (predictor variable, x) (West & Post 2016). To fit a model, the empiricist provides candidate models (e.g., Ricker Model, Beverton-

Holt model, a new model) that each propose a different relationship between x and y (e.g., straight line, hump-shaped), along with the data for the variables in the model (e.g., x and y). Then statistical methods such as maximum likelihood are used to estimate the attributes (parameters) of that relationship (e.g., the y intercept, the slope, and the curvature) of the model with the highest probability of explaining the observed data. These statistical methods usually quantify the deviation between the model expectation and the data, and then iteratively alter the parameters to make the deviations smaller until the best-fitting model is found (Bolker 2008). The outputs of model-fitting and model comparison are estimates of the parameters, their uncertainties (errors), and relative support for the different candidate models. The empiricist may be interested in these parameter values themselves (e.g., determining the space clearance rate of a given predator on a specific prey), or how they change with an experimental manipulation (e.g., determining how space clearance rates vary with temperature, as in DeLong and Lyon (2020)).

A second way to use equations in empirical research is to collect data that can simply be plugged into an equation as predictor variables in order to calculate the response variable. In this case, the collected data are the variables and parameters on the right-hand side of the equation, and the variable on the left-hand side of the equation is the value of interest that is calculated. Examples of this approach include estimating maximum sustainable yield in fisheries from population growth rate and carrying capacity (Schaefer 1954) and estimating the spread rate of invading organisms from dispersal and population growth rates (Andow et al. 1990; Kolmogorov 1937). To do this, the empiricist collects the necessary data and enters them as predictor variables into the equation. The estimated outputs that are generated can then be used in subsequent analyses; for example, determining how species spread rates are affected by habitat

type (Andow et al. 1990). One issue to be aware of with this approach is that collected data will have error surrounding them, and determining the proper way to carry this error through to the final estimated data can require careful consideration (e.g., Terry et al. (2021)).

Approach 4. Test the model assumptions

A final way that empiricists can use theory is to directly test the assumptions that underlie theoretical models (Fig. 1, Box 2). This type of empirical research has a strong link to theory, and is often conducted with the express purpose of informing future theoretical work (Price *et al.* 2012; Servedio *et al.* 2014). This type of research has the potential to make a major impact in the field, particularly when the tested assumptions are widely invoked. For example, empirical research that challenged the widespread assumptions that evolution is slower than ecology (Losos et al. 1997; Reznick et al. 1990) and that interspecific niche differences play a major role in structuring ecological communities (Hubbell 2001) have inspired major bodies of research. Looking forward, many of the simplifying assumptions of The Metabolic Theory of Ecology have not been adequately verified by empirical research (Price et al. 2012), and the widely adopted assumption that pairwise interactions can adequately capture community dynamics, which underlies much of coexistence research, remains largely untested (Levine *et al.* 2017).

How to do it

The first step in testing model assumptions is to identify assumptions that are likely to have significant impacts on the conclusions drawn in a field but are not yet well-supported by empirical research. This can be achieved through a broad reading of both the theoretical and empirical work in the field, and widely applied assumptions may be particularly likely to come to light during a literature review or meta-analysis. Assumptions that have substantial effects on model outcomes will hopefully be identified explicitly in theoretical papers (see Box 2 “Meeting

model assumptions”) and can also be explored using the techniques for understanding equations provided below (particularly Tip #4 “Achieve a working understanding of equations”). The next step is to collect the data required to support or refute the assumption. For example, while classic foraging theory assumes that per capita predator feeding rates are independent of predator density (Holling 1966), empirical data on feeding rates across predator densities have refuted this assumption and prompted revised models (DeLong and Vasseur 2011; Novak et al. 2017; Vucetich et al. 2002).

One consideration when designing this type of experiment is how broadly the results can be applied. If the goal is to assess whether the assumption is met in the type of system to which this theory is most often or most appropriately applied, then testing in a representative system makes sense (e.g., Bernhardt et al. (2018)). If the goal is to make a broad statement about whether an assumption is met, in general, then data from multiple systems may be needed (e.g., DeLong and Vasseur (2011)). Ideally, the final step for empirical research that tests model assumptions will be that the results of such experiments will feed back to inform and improve subsequent theoretical work.

A toolkit for understanding equations

So far, we have described how theory is created and have highlighted several ways that empiricists can integrate theory into their research to attain a better understanding of nature. However, many theories in ecology are expressed in the language of math, and in order to effectively use mathematical theory, empiricists must be able to understand it. A helpful first step towards understanding mathematical theory is to determine generally what type of model one is dealing with by answering some basic questions (Box 1). The next step is likely the biggest hurdle most empiricists face: understanding equations. We tackle this challenge by providing

concrete tips for visualizing functions, interpreting common symbols and figures, and deciphering their biological meanings. For readers looking for a more comprehensive treatment of these topics, we also provide a list of textbook sections that cover these topics in much greater depth (Table S1).

To anchor our discussion, we begin by presenting a modified Rosenzweig-MacArthur consumer-resource model where both the consumer (with density represented by the variable C) and the resource (with density R) grow and impact each other (Rosenzweig and MacArthur 1963):

$$\frac{dR}{dt} = \underbrace{rR}_{\text{parameter}} \left(\underbrace{1 - \frac{R}{K}}_{\text{expression}} \right) - \underbrace{\frac{aR}{1 + ahR}C}_{\text{term}} \quad (1.1)$$

$$\frac{dC}{dt} = e \frac{aR}{1 + ahR} C - mC \quad (1.2)$$

Now, did you read the equations or skip past them? Our first piece of advice is to give the math a chance: spend five minutes trying to get to know it. The more time one spends with equations, the more familiar reading new equations becomes. Try to not get bogged down by new symbols (e.g., Greek lettering), as often the specific choice of symbols used is arbitrary or based on precedent. Here, C and R are the densities of individuals in each population, and dC/dt and dR/dt are rates of change of each population's size; note that dC/dt represents a rate of change in continuous-time models, whereas discrete-time models are expressed as $C_{t+1} = f(C_t)$ which represents the value of a variable at the next time step (Box 1). When $dC/dt > 0$, the value of C is

increasing over time, whereas C is decreasing when $dC/dt < 0$. In eqn. 1.1, resources experience some maximum per capita growth rate in the absence of competition or consumers, denoted by r , and have carrying capacity K . In eqn. 1.2, a is the space clearance rate (a.k.a attack rate; the area or volume cleared of prey per predator per prey unit time), h is the amount of time a consumer spends handling each prey item (e.g., time to kill, eat, digest, etc.) that would otherwise be spent searching for prey, e is the conversion rate of consumed resources into new consumer individuals, and m is the per capita rate of consumer death.

What's the best way to begin approaching these equations? We start by unpacking an equation's basic anatomy and considering the terminology that can be used to discuss different elements of equations. Each symbol in an equation can be classified as either a variable, a parameter, or an operator. As described above ('How is mathematical theory created?'), a **variable** is a quantity whose value changes either dependently (e.g., population size) or independently (e.g., time) of other variables. **Parameters** are fixed quantities that remain constant within a given equation (e.g., consumer death rate). **Constants** are numbers that enter equations according to some assumption made by the theoretician. For example, in eqn. 1.2, the one in $1-R/K$ indicates that resources grow at 100% of their intrinsic growth rate if R is near zero. Sometimes constants are called **coefficients** or **constant factors** when multiplying a variable (e.g., the coefficient rN in an equation of exponential growth represents the fact that every individual will produce r surviving offspring on average). **Operators** describe how different quantities in equations interact with each other, including simple operations from algebra (i.e., addition, multiplication, etc.) and complex ones, such as those associated with set theory (e.g., \cap) or calculus (e.g., \int ; Fig. 2).

Collections of symbols also have their own terminology. A **term** is a set of variables and/or parameters that are multiplied or divided together or placed inside parentheses. There are two terms on the right-hand side of eqn. 1.2: $e(aRC)/(1+ahR)$ and mC . These terms may be composed of multiple **factors** that multiply together to give the term (e.g., m times C). An **expression** is the full set of terms contained on the same side of the equal sign. Any one equation can be rearranged, or “expressed,” in different ways, resulting in different expressions that maintain the equivalence between the left-hand side and right-hand side of the equation (sometimes abbreviated to ‘LHS’ and ‘RHS’ in theory papers, respectively). Clearly, there is a specific terminology associated with different elements of equations that should not be used interchangeably (i.e., variable \neq parameter, term \neq expression).

Now, let’s get into some concrete tips.

Tip 1: Align the math with the biology. Math provides an abstraction of nature, but back-translating math into biology can help make equations interpretable and can make it easier to identify the model’s assumptions. In the equations above, notice that the population growth of resource R depends on the population size of consumers C and vice versa (i.e., R and C appear in both equations). This interdependence makes intuitive biological sense because the consumer population relies on resources to produce offspring, and as consumers consume resources, the resource population is depleted. Also notice that in these equations, the resource can exist without the consumer (if $C=0$ eqn. 1.1, the growth of R reverts to logistic growth), but the consumer cannot persist without the resource (if $R=0$ and $C>0$ initially in eqn. 1.2, C will decline to extinction (i.e., $C=0$) over time because $dC/dt < 0$). Again, this intuitively makes sense (see Tip 4, for tips on how to understand any equation) and it illustrates how models often build upon previous models (here adding a consumer to a model of resource growth). Notice also that the

equation for resource population growth includes a carrying capacity (K), which causes population growth to slow as the value of R (resource density) approaches K (i.e., R/K approaches 1, causing the quantity in the parentheses to approach 0). This indicates that, in this particular set of consumer-resource equations, the resource R is biological and is itself limited by density-dependent competition. This equation for R thus fits the behavior of a plant but not an abiotic resource that enters and exits the system in a density-independent fashion (i.e., sunlight). Try picturing two organisms that adhere to the biological assumptions above in order to make the abstract math more tangible. Lastly, if the consumer spends relatively little time handling a resource relative to the time it takes to find and attack that resource (so $h \approx 0$, for example, a sessile filter feeder), then the rate of resource consumption by each consumer increases linearly with the population size of the resource (because the denominator in eqn. 1.2 equals 1). If, by contrast, substantial time is spent handling resources (i.e., $h \gg 0$, for example, carnivores that hunt large prey), then the rate of resource consumption saturates when more resources are present than the consumer has time to consume.

Tip 2: Think in terms of stocks and flows. On a basic level, equations describe how variables grow or shrink in response to other variables. The **stock** is the amount of a given variable that exists at any point in time (e.g., population size) whereas **flows** are inflows (i.e., individuals added by birth) and outflows (i.e., individuals being removed by death) that affect the amount of stock and its rate of change. It can help to keep track of which parameters represent inflows versus outflows. Drawing out a **stock and flow diagram** (Box Fig. 1a) by linking stocks (typically represented by circles) and flows (arrows) can be helpful, especially when multiple terms cause any one variable to shrink or grow or when multiple variables are linked. Although the conventions for stock and flow diagrams vary across the literature, we use a recommended

approach that maximizes connections between the equations and the diagram (Otto and Day 2007) (see also Ogbunugafor and Robinson (2016)).

Tip 3: Verbally summarize complex terms. Some mathematical terms are complex and involve many symbols. For example, the term $(eaRC/(I+ahR))$ is the consumer birth rate, but it can be broken into three main parts: the number of consumers (C), the foraging rate per consumer $(aR/(I+ahR))$, and the number of consumers produced per resource eaten (e). By breaking down multi-symbol terms in this way, we can see which mechanisms are being invoked (here, individual-level foraging, satiation, reproduction) to generate higher-order phenomena (here, population-level growth). Thus, although complex, we can simplify eqn. 1.2 to read as just two parts: individuals added via birth due to consumption minus individuals lost via death. These simplifications can be written out in words as we have done in our description of $eaRC/(I+ahR)$, or can be shown by annotating the equations to reflect the summarized descriptions. Doing so reduces the number of specific symbols that need to be tracked. Note that some parameters in equations may themselves be a simplification made by the author. For example, parameter h in equation 1.2 is a consumer's handling time, but h can be expanded further to include time spent subduing a resource, time spent consuming it, and time spent resting following consumption. Rearranging equations to build new expressions (while maintaining left-hand side and right-hand side equivalence) can produce more biologically intuitive descriptions of terms (e.g., dividing both sides of equation 1.2 by C in order for the left-hand side to be expressed as a per capita rate).

Tip 4: Achieve a working understanding of equations. Well enough is better than not at all, but what level of understanding is sufficient to be useful? For example, one may not need to know at what exact value of a variable a function reaches an inflection point, but understanding

whether the function is saturating or non-saturating is still useful. One trick is to think in terms of what happens to a dependent variable when the independent variable takes on a particular value, especially extreme values. For example, consider what happens in an equation if a variable is zero vs a large number: perhaps an asymptote is reached, or the function may become undefined (e.g., division by 0), meaning that the value chosen is outside the **domain** (set of possible values) of the function. Solving equations with arbitrary values of variables also can help determine whether a function is linear or nonlinear, or if relationships between variables are negative or positive. This is a “plug and play” approach where plug = input value and play = calculate output value. This approach can be used to understand the relationship between two variables for some specific combination of parameter values, or to explore how different parameter values change the relationship between variables.

Tip 5: Visualize functions precisely by plotting them. Tip 4 described how to generally understand what a function looks like. For some purposes, though, this is not sufficient. Although a theoretician’s ideal solution to this problem would be to find analytical solutions, doing so may require expert knowledge and may not be feasible for some complex equations. For someone trying to better understand a theory, another option is to input equations in a spreadsheet (e.g., Excel) or a programming application (e.g., R) and use plotting to explore the consequences of changing different parameter values or the initial value of variables on the dynamics of the state variable itself (e.g., N_t). This is essentially a simple numerical analysis. A worked example of how to explore equations using simple numerical analysis can be found in Chapter 3 of Bolker (2008). Alternatively, interactive applications for exploring certain models can be found online (e.g., Shiny apps, see Kandlikar (2021)).

Tip 6: Identify common signposts. In reading theory papers, one might notice a range of symbols of various script types (e.g., Roman, Greek) and typefacing (e.g., bolding, capitalization, cursive script). Although any one symbol may be used to denote entirely different variables in different papers, or even worse, different symbols may be used to denote the same biological process in different papers entirely at the discretion of the author, some symbols are used consistently and have a specific meaning (Edwards and Auger- Méthé 2019). When these types of symbols appear in theory papers to symbolize complex operations or data structures, they are rarely defined because they are considered common knowledge. Not knowing what these symbols represent, or even which symbols are specific to that paper and which hold some generally accepted meaning, can be a major barrier to understanding equations. In Figure 2, we present a list of common symbols/typefaces, describe what they mean mathematically, and provide an example of their use in a biological context. Although we cannot cover every symbol that will be encountered in theory papers, simply knowing that undefined symbols and specific typefaces often hold specific meaning should help you to recognize when and what to ask of a search engine (most symbols can be copied and pasted into search engines). We also cover some commonly encountered mathematical objects and operations (e.g., Jacobian matrices) in Figure 2.

Tip 7: Consider the fact that some information might be missing, presented unclearly, or incorrect. In some instances, confusion on the part of the reader may stem from the presentation of the material itself (e.g., undefined terms, unclear units) rather than the reader's knowledge. We hope that our discussion of some of the barriers to empiricists understanding theory will encourage theoreticians to present their work in a way that can be more easily interpreted by a non-expert. And while we have focused our discussion on tests of theory that do not involve

determining whether or not the math is correct, it may also be helpful for empiricists to keep in mind that theory is not infallible and that in some instances the math may indeed be incorrect.

We suggest that empiricists, after putting in some effort to understand a theory paper and arriving at unresolved uncertainties, contact the authors for clarification.

Conclusion

We hope that by demystifying theoretical work, the perspectives presented here will help break down current barriers to the integration of, and feedback between, theoretical and empirical research in ecology. We particularly hope that this paper is valuable to early career empiricists starting out in the field. Charles Elton, when reviewing Alfred Lotka's book on the mathematics of populations, famously critiqued, "like most mathematicians, [Lotka] takes the hopeful biologist to the edge of the pond, points out that a good swim will help [their] work, and then pushes [them] in and leaves [them] to drown" (Elton 1935). Unfortunately, nearly a century later, hopeful empiricists are still regularly being left to flail around in the murky waters of theory. Here we offer a life raft and encourage empiricists to dip into the shallow waters of theory and to slowly gain the confidence, capability, and curiosity to venture deeper.

Acknowledgements

We are grateful to the Levine Lab, the Germain Lab, and the O'Connor Lab for helpful discussions about the ideas presented in this paper. We also thank the Downtowner Deli in Louisville for initially bringing this working group together. Funding was provided by a UBC Grant for Catalyzing Research Clusters to TNG and RMG, an NSERC Postdoctoral Fellowship and a Banting Postdoctoral Fellowship to TNG, and an NSERC Discovery Grant to RMG (2019-04872).

Author contributions

TNG and RMG conceived of the idea for the paper and wrote the manuscript, AS, PJK and MAB helped make the figures and boxes, and all authors provided feedback and revisions on manuscript drafts.

Data accessibility

Annotated code for Box Figure 1 is publicly available on GitHub (<https://github.com/mabarbour/foodweb-theory/releases/tag/v2.0>), and we also provide a more user-friendly website to view the code (<https://mabarbour.github.io/foodweb-theory/rosenzweig-macarthur.html>). This code has also been archived with Zenodo (<https://doi.org/10.5281/zenodo.5154139>).

Literature Cited

- Andow, D., P. Kareiva, S. A. Levin, and A. Okubo. 1990. Spread of invading organisms. *Landscape ecology* 4:177-188.
- Bernhardt, J. R., J. M. Sunday, and M. I. O'Connor. 2018. Metabolic theory and the temperature-size rule explain the temperature dependence of population carrying capacity. *American Naturalist* 192:687-697.
- Bodner, K., C. Brimacombe, E. Chenery, A. Greiner, A. McLeod, S. Penk, and J. S. V. Soto. 2020. Ten simple rules for tackling your first mathematical models: A guide for graduate students by graduate students. *PLoS Computational Biology* 17:e1008539.
- Bolker, B. M. 2008, *Ecological models and data in R*, Princeton University Press.
- Britton, T., F. Ball, and P. Trapman. 2020. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science* 369:846-849.
- Broekman, M. J., H. C. Muller-Landau, M. D. Visser, E. Jongejans, S. Wright, and H. de Kroon. 2019. Signs of stabilisation and stable coexistence. *Ecology letters* 22:1957-1975.
- Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. *Ecology* 85:1771-1789.
- Carroll, I. T., B. J. Cardinale, and R. M. Nisbet. 2011. Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology* 92:1157-1165.
- Caswell, H. 1988. Theory and models in ecology: a different perspective. *Ecological modelling* 43:33-44.
- Charnov, E. L. 1976. Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9:129-136.
- Chase, J. M., A. A. Burgett, and E. G. Biro. 2010. Habitat isolation moderates the strength of top-down control in experimental pond food webs. *Ecology* 91:637-643.

- Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31:343-366.
- Clark, J. S., C. Fastie, G. Hurtt, S. T. Jackson, C. Johnson, G. A. King, M. Lewis et al. 1998. Reid's paradox of rapid plant migration: dispersal theory and interpretation of paleoecological records. *BioScience* 48:13-24.
- Cohen, D. 1966. Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology* 12:119-129.
- Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees, Pages 298-312 *in* P. J. D. Boer, and G. R. Gradwell, eds. *Dynamics of populations: Proceedings of the Advanced Study Institute on Dynamics of numbers in populations*, Oosterbeek, the Netherlands, 7-18 September 1970.
- Darwin, C. 1859, *On the origins of species by means of natural selection*. London, Murray.
- Dee-Lucas, D., and J. H. Larkin. 1991. Equations in scientific proofs: Effects on comprehension. *American Educational Research Journal* 28:661-682.
- DeLong, J. P., and S. Lyon. 2020. Temperature alters the shape of predator–prey cycles through effects on underlying mechanisms. *PeerJ* 8:e9377.
- DeLong, J. P., and D. A. Vasseur. 2011. Mutual interference is common and mostly intermediate in magnitude. *BMC ecology* 11:1.
- Dennis, B., R. A. Desharnais, J. Cushing, and R. Costantino. 1995. Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments. *Ecological Monographs* 65:261-282.
- Duffy, M. A., S. R. Hall, A. J. Tessier, and M. Huebner. 2005. Selective predators and their parasitized prey: are epidemics in zooplankton under top- down control? *Limnology and Oceanography* 50:412-420.

- Edwards, A. M., and M. Auger- Méthé. 2019. Some guidance on using mathematical notation in ecology. *Methods in Ecology and Evolution* 10:92-99.
- Elton, C. 1935. Eppur Si Muove. *Journal of Animal Ecology* 4:148-150.
- Fawcett, T. W., and A. D. Higginson. 2012. Heavy use of equations impedes communication among biologists. *Proceedings of the National Academy of Sciences* 109:11735-11739.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1-15.
- Ferrier, S., K. Ninan, P. Leadley, R. Alkemade, L. Acosta, H. Akçakaya, L. Brotons et al. 2016. The methodological assessment report on scenarios and models of biodiversity and ecosystem services. Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services: Bonn, Germany.
- Gause, G. 1935. Experimental demonstration of Volterra's periodic oscillations in the numbers of animals. *Journal of Experimental Biology* 12:44-48.
- Godoy, O., and J. M. Levine. 2014. Phenology effects on invasion success: insights from coupling field experiments to coexistence theory. *Ecology* 95:726-736.
- Grainger, T. N., and B. Gilbert. 2016. Dispersal and diversity in experimental metacommunities: linking theory and practice. *Oikos* 125:1213-1223.
- Grime, J. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *American naturalist*:1169-1194.
- Grimm, V. 1994. Mathematical models and understanding in ecology. *Ecological modelling* 75:641-651.
- Haller, B. C. 2014. Theoretical and empirical perspectives in ecology and evolution: a survey. *BioScience* 64:907-916.
- Hardin, G. 1960. The competitive exclusion principle. *Science* 131:1292-1297.

- Harms, K. E., S. J. Wright, O. Calderón, A. Hernandez, and E. A. Herre. 2000. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* 404:493-495.
- Harte, J., and E. A. Newman. 2014. Maximum information entropy: a foundation for ecological theory. *Trends in ecology & evolution* 29:384-389.
- Hastings, A. 2001. Transient dynamics and persistence of ecological systems. *Ecology Letters* 4:215-220.
- Holling, C. S. 1966. The functional response of invertebrate predators to prey density. *The Memoirs of the Entomological Society of Canada* 98:5-86.
- Hubbell, S. P. 2001, *The unified neutral theory of biodiversity and biogeography*. Princeton, Princeton University Press.
- Huffaker, C. 1958. Experimental studies on predation: dispersion factors and predator-prey oscillations. *Hilgardia* 27:343-383.
- Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* 104:501-528.
- Jost, L. 2006. Entropy and diversity. *Oikos* 113:363-375.
- Kandlikar, G. S. 2021. Lotka Volterra Competition Shiny App.
<https://gauravsk.shinyapps.io/lotka/>.
- Kareiva, P. 1989. Renewing the dialogue between theory and experiments in population ecology, Pages 68-88 in R. M. May, J. Roughgarden, and S. A. Levin, eds. *Perspectives in Ecological Theory*, Princeton University Press.
- Ke, P. J., and J. Wan. 2020. Effects of soil microbes on plant competition: a perspective from modern coexistence theory. *Ecological Monographs* 90:e01391.
- Kneitel, J. M., and T. E. Miller. 2003. Dispersal rates affect species composition in metacommunities of *Sarracenia purpurea* inquilines. *American Naturalist* 162:165-171.

- Knight, J. 2003. Null and void. *Nature* 422:554–555.
- Kokko, H. 2007, *Modelling for field biologists and other interesting people*, Cambridge University Press.
- Kolmogorov, A. N. 1937. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. Moskow, Ser. Internat., Sec. A* 1:1-25.
- Krebs, C. J. 1988. The experimental approach to rodent population dynamics. *Oikos*:143-149.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution*:1210-1226.
- Laubmeier, A., K. Wootton, J. Banks, R. Bommarco, A. Curtsdotter, T. Jonsson, T. Roslin et al. 2018. From theory to experimental design—Quantifying a trait-based theory of predator-prey dynamics. *PLoS One* 13:e0195919.
- Laubmeier, A. N., B. Cazelles, K. Cuddington, K. D. Erickson, M.-J. Fortin, K. Ogle, C. K. Winkle et al. 2020. *Ecological Dynamics: Integrating Empirical, Statistical, and Analytical Methods*. *Trends in Ecology & Evolution*.
- Levine, J. M., and M. Rees. 2004. Effects of temporal variability on rare plant persistence in annual systems. *American Naturalist* 164:350-363.
- Levins, R. 1966. The strategy of model building in population biology. *American scientist* 54:421-431.
- Loehle, C. 1987. Hypothesis testing in ecology: psychological aspects and the importance of theory maturation. *Quarterly Review of Biology* 62:397-409.
- Łomnicki, A. 1988. The place of modelling in ecology. *Oikos*:139-142.
- Losos, J. B., K. I. Warheit, and T. W. Schoener. 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature* 387:70-73.

- Lotka, A. J. 1926. Elements of physical biology. *Science Progress in the Twentieth Century* (1919-1933) 21:341-343.
- MacArthur, R. 1970. Species packing and competitive equilibrium for many species. *Theoretical population biology* 1:1-11.
- MacArthur, R. H., and E. O. Wilson. 1967, *The theory of island biogeography*, v. 1, Princeton University Press.
- Marquet, P. A., A. P. Allen, J. H. Brown, J. A. Dunne, B. J. Enquist, J. F. Gillooly, P. A. Gowaty et al. 2014. On theory in ecology. *BioScience* 64:701-710.
- Moeller, H. V., E. Peltomaa, M. D. Johnson, and M. G. Neubert. 2016. Acquired phototrophy stabilises coexistence and shapes intrinsic dynamics of an intraguild predator and its prey. *Ecology letters* 19:393-402.
- Mouquet, N., and M. Loreau. 2003. Community patterns in source- sink metacommunities. *American Naturalist* 162:544-557.
- Narwani, A., M. A. Alexandrou, T. H. Oakley, I. T. Carroll, and B. J. Cardinale. 2013. Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters* 16:1373-1381.
- Novak, M., C. Wolf, K. E. Coblenz, and I. D. Shepard. 2017. Quantifying predator dependence in the functional response of generalist predators. *Ecology Letters* 20:761-769.
- O'Connor, M. I., B. Gilbert, and C. J. Brown. 2011. Theoretical predictions for how temperature affects the dynamics of interacting herbivores and plants. *American Naturalist* 178:626-638.
- Ogbunugafor, C. B., and S. P. Robinson. 2016. OFF1 models: novel schema for dynamical modeling of biological systems. *PloS one* 11:e0156844.
- Okubo, A. 1986. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Advances in Biophysics* 22:1-94.

- Okubo, A., and S. A. Levin. 2013, Diffusion and ecological problems: modern perspectives, v. 14, Springer Science & Business Media.
- Otto, S. P., and T. Day. 2007, A biologist's guide to mathematical modeling in ecology and evolution, Princeton University Press.
- Otto, S. P., and A. Rosales. 2020. Theory in service of narratives in evolution and ecology. *American Naturalist* 195:290-299.
- Parain, E. C., S. M. Gray, and L.-F. Bersier. 2019. The effects of temperature and dispersal on species diversity in natural microbial metacommunities. *Scientific Reports* 9:1-10.
- Phillips, R. 2015. Theory in biology: Figure 1 or Figure 7? *Trends in cell biology* 25:723-729.
- Pianka, E. R. 1970. On r-and K-selection. *American naturalist* 104:592-597.
- Pickett, S. T., J. Kolasa, and C. G. Jones. 2010, Ecological understanding: the nature of theory and the theory of nature, Elsevier.
- Pielou, E. C. 1969, An introduction to mathematical ecology, Wiley Interscience.
- Platt, J. R. 1964. Strong inference. *Science* 146:347-353.
- Price, C. A., J. S. Weitz, V. M. Savage, J. Stegen, A. Clarke, D. A. Coomes, P. S. Dodds et al. 2012. Testing the metabolic theory of ecology. *Ecology letters* 15:1465-1474.
- Prugh, L. R., K. E. Hodges, A. R. Sinclair, and J. S. Brashares. 2008. Effect of habitat area and isolation on fragmented animal populations. *Proceedings of the National Academy of Sciences* 105:20770-20775.
- Reznick, D. A., H. Bryga, and J. A. Endler. 1990. Experimentally induced life-history evolution in a natural population. *Nature* 346:357-359.
- Rosenzweig, M. L., and R. H. MacArthur. 1963. Graphical representation and stability conditions of predator-prey interactions. *American Naturalist* 97:209-223.

- Rossberg, A. G., G. Barabás, H. P. Possingham, M. Pascual, P. A. Marquet, C. Hui, M. R. Evans et al. 2019. Let's train more theoretical ecologists – here is why. *Trends in Ecology & Evolution* 34:759-762.
- Roughgarden, J. 1971. Density- dependent natural selection. *Ecology* 52:453-468.
- Rutledge, R. W., B. L. Basore, and R. J. Mulholland. 1976. Ecological stability: an information theory viewpoint. *Journal of Theoretical Biology* 57:355-371.
- Scheiner, S. M. 2013. The ecological literature, an idea- free distribution. *Ecology Letters* 16:1421-1423.
- Servedio, M. R. 2020. An effective mutualism? The role of theoretical studies in ecology and evolution. *American Naturalist* 195:284-289.
- Servedio, M. R., Y. Brandvain, S. Dhole, C. L. Fitzpatrick, E. E. Goldberg, C. A. Stern, J. Van Cleve et al. 2014. Not just a theory—the utility of mathematical models in evolutionary biology. *PLoS Biol* 12:e1002017.
- Shoemaker, L. G., J. A. Walter, L. A. Gherardi, M. H. DeSiervo, and N. I. Wisnoski. 2021. Writing mathematical ecology: A guide for authors and readers. *Ecosphere* 12:e03701.
- Simberloff, D. S., and E. O. Wilson. 1969. Experimental zoogeography of islands: the colonization of empty islands. *Ecology* 50:278-296.
- Stump, S. M., and L. S. Comita. 2018. Interspecific variation in conspecific negative density dependence can make species less likely to coexist. *Ecology letters* 21:1541-1551.
- Terry, J. C. D., J. Chen, and O. T. Lewis. 2021. Natural enemies have inconsistent impacts on the coexistence of competing species. *Journal of Animal Ecology*.
- Tilman, D. 1977. Resource competition between plankton algae: an experimental and theoretical approach. *Ecology* 58:338-348.
- Uszko, W., S. Diehl, and J. Wickman. 2020. Fitting functional response surfaces to data: a best practice guide. *Ecosphere* 11:e03051.

- Venable, D. L. 2007. Bet hedging in a guild of desert annuals. *Ecology* 88:1086-1090.
- Volterra, V. 1927, *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*, C. Ferrari.
- Vucetich, J. A., R. O. Peterson, and C. L. Schaefer. 2002. The effect of prey and predator densities on wolf predation. *Ecology* 83:3003-3013.
- West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. *Science* 276:122-126.
- White, C. R., and D. J. Marshall. 2019. Should We Care If Models Are Phenomenological or Mechanistic? *Trends in Ecology & Evolution* 34:276-278.

References cited only in online supporting information

- Bodine, E.N., S. Lenhart, & L. J. Gross. 2014. *Mathematics for the life sciences*. Princeton University Press.
- Case, T. J. 2000. *Illustrated guide to theoretical ecology*. Oxford University Press.
- Caswell, H. 2006. *Matrix population models: construction, analysis, and interpretation*. Second edn. Sinauer Associates.
- Crawley, M. J. 2012. *The R book*. Second edn. John Wiley & Sons.
- Gotelli, N. J. 2008. *A primer of ecology*. Fourth edn. Sinauer Associates Sunderland, Massachusetts, USA.
- Hilborn, R., & M. Mangel. 1997. *The ecological detective: confronting models with data*. Princeton University Press.
- McCann, K. S. 2011. *Food webs (MPB-50)*. Princeton University Press.
- McElreath, R. 2020. *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press.

Pastor, J. 2008. Mathematical ecology of populations and ecosystems. John Wiley & Sons.

Roughgarden, J. 1998. Primer of ecological theory. Pearson.

Box 1. Types of mathematical models

Determining the general attributes of a theoretical model by asking the following three questions can help the empiricist understand a model and determine whether or not a model is a good fit for their research.

How does the model deal with the passage of time?

Determining whether, and how, a model involves the passage of time will provide insight into the type of study system it describes. Models that describe how a system changes through time are called dynamical models and are very common in ecology. Some dynamical models include discrete time steps (e.g., from one life stage to the next or one generation to the next), whereas others model time continuously. **Discrete-time models** will have something like $n(t + 1)$ or n_{t+1} on the left-hand side of the equation (which represents the value of a variable at the next time step), whereas **continuous-time models** will have something like dn/dt on the left-hand side of the equation (which represents a *change* in the value of a variable over time). This distinction can be important when determining whether a model matches a given study system. For example, because population sizes change when organisms reproduce or die, if the focal organism reproduces in discrete time steps with non-overlapping generations (e.g., annual plants, birds), then a discrete-time model may be most applicable. However, if the organism reproduces continuously (e.g., zooplankton, aphids), a continuous-time model might make more sense. While understanding how time is passing in a model can be helpful to the empiricist, note that theoreticians may model time in a certain way for mathematical convenience and that it is possible to convert a model between discrete and continuous time under the appropriate assumptions (Otto and Day 2007). Finally, some models do not include the passage of time at all,

including most metabolic theory (Brown et al. 2004) and optimal foraging theory (Charnov 1976).

Is the model mechanistic or phenomenological?

Although the distinction between mechanistic and phenomenological models is often fuzzy and will depend on the perspective of the investigator and the theoretical framework in which they are working, thinking about this distinction can help the empiricist decide whether a model fits their needs (Otto and Day 2007; White and Marshall 2019). **Mechanistic models** explicitly include underlying biological mechanisms, or put another way, explicitly model dynamics happening at a different level of organization. For example, in models that describe resource competition, a mechanistic model will model the amount of each resource produced and consumed through time. Classic examples include MacArthur's consumer-resource model (MacArthur 1970) and Tilman's resource competition model (Tilman 1977). In both of these, the resources being competed for are included in the model. In contrast, **phenomenological models** are focused on the patterns (i.e., the phenomena) that emerge from an underlying process, rather than the process itself. For example, a phenomenological model of populations competing for a shared resource tracks population dynamics of consumers without explicitly including the underlying mechanisms that drive those responses (Otto and Day 2007; White and Marshall 2019). Instead of including the entity being consumed in the model, competition for limited resources is implied by including terms such as the carrying capacity or competition coefficients. A classic example is the Lotka-Volterra competition model (Lotka 1926; Volterra 1927).

If the empiricist has an interest in the underlying mechanisms (e.g., the nature of resource competition) and has the required information to apply a mechanistic model (e.g., the key resources underlying competition are known), then a mechanistic model can offer deeper insights

about the system. If they do not, then a phenomenological model may be more appropriate and does not require as in-depth an understanding of underlying drivers.

Are the predictions qualitative or quantitative?

Qualitative predictions state that x is related to y in some way and are very common in ecological theory. For example, The Theory of Island Biogeography predicts that species richness is positively related to island size and proximity to the mainland (MacArthur and Wilson 1967). A qualitative prediction may describe a broad pattern such as “species reach carrying capacity” or x affects y , or may be more specific, such as “carrying capacity scales linearly with intrinsic growth rate” or “there is a hump-shaped relationship between x and y .” By contrast, **quantitative predictions** include specific numerical values and are more common in ecosystem theory and metabolic theory. For example, metabolic theory predicts that an organism’s metabolic rate scales with the $3/4$ power of its body mass (West et al. 1997). Considering the level of specificity of a model’s predictions can help the empiricist start to think more precisely about the possible outcomes of an experiment that would support a given prediction.

Box 2. Meeting model assumptions

A critical step in designing empirical research that provides a robust test of theoretical predictions is to ensure that the study meets as many of the model's assumptions as possible. This is not always straightforward, as model assumptions may not be explicitly stated in words in a theory paper. If assumptions are stated explicitly, they will often be found in the “model” section of the paper, and some indication will hopefully be given of how critical the different assumptions are for the predictions generated by the model. For example, there may be a paragraph in the discussion that mentions the assumptions and limitations of the model and is somewhat analogous to a “caveats” section in an empirical paper. If this is not in the paper, the tips provided in ‘A toolkit for understanding equations’ may help empiricists uncover assumptions that are described mathematically but not verbally and feel empowered to explore how sensitive the outcome is to the assumed relationship between variables and parameters using a “plug and play” approach (Tip 4).

Some common assumptions in models used in ecological theory are:

- the system is analyzed or observed at dynamic equilibrium
- species interactions are competitive (i.e., not facilitative)
- resources are limited
- individuals interact with all other individuals with equal likelihood (i.e., no spatial structure)
- individuals within a species are identical in terms of how they interact, reproduce, and survive (i.e., no intraspecific variation)
- there is a time-scale separation between ecology (fast) and evolution (slow)

- there is a time-scale separation between resource growth (fast) and consumption (slow)
- dispersal rates are homogenous across species and are not spatially explicit (i.e., ‘global’ dispersal)
- the outcomes of biological processes are deterministic (i.e., no stochasticity)
- population growth is density-dependent and follows a simple functional form (e.g., linearly declining per capita growth rate as the population size increases in the logistic equation)
- predator consumption rates saturate as prey densities increase (i.e., ‘Type II functional response’)

Note that some of these assumptions are made for mathematical convenience (e.g., no spatial structure, no stochasticity) whereas others reflect biological realities (e.g., resources are limited, population growth is density-dependent). Likewise, some of these assumptions should be evident from the mathematical equations themselves (e.g., species interactions are competitive, predator consumption rates saturate as prey density increases), whereas others must be stated explicitly in words in order to be communicated (e.g., no spatial structure, no intraspecific variation).

While keeping these types of distinctions in mind can help the empiricist navigate the task of determining what a model’s assumptions are and how closely they need to be matched experimentally, deciding how perfectly assumptions need to be met in order for a model to be useful is a subjective exercise. Our stance is that empirical work does not need to perfectly satisfy every model assumption and that the focus should be on understanding which assumptions are critical to the outcomes and which are more flexible. Ideally this is outlined, or at least alluded to, in the theory paper (Servedio 2020). Worthwhile questions to ask include: If

these assumptions were broken, how would this change the predicted outcome? And: is it possible to meet the critical assumptions of this model in the biological system of interest, or is a different model or system needed?

Thinking about assumptions can be a useful philosophical exercise for empiricists at both the microscale of an individual experiment and the macroscale of the entire field. At the microscale, there are assumptions in every empirical study, even those not based on theory, and grounding experiments strongly in theory can help bring assumptions to light and force researchers to explicitly state what they are and when they are bent or broken. This can encourage the empiricist to think more deeply about experimental design and can make it easier for future researchers to build on their work. At the macroscale, it is useful for the discipline as a whole to regularly examine the common assumptions that have become the default and to consider whether they indeed reflect reality, and how outcomes might change if these assumptions were to be broken (see ‘Test the Model Assumptions’).

Box 3. How to interpret common figures in ecological theory papers

Here, we present five common types of figures in ecological theory (Box Fig. 1) and walk through their interpretation. Each panel shows a different way of representing the structure (panel a) or the outcome of model simulations (panels b-e) of the same model (eqn. 1 in the main text).

Stock and flow diagram (Box Fig. 1a): These diagrams visually summarize a model by showing its components and how they interact. Box Fig. 1a shows one of many ways to draw a stock and flow diagram and follows conventions from Otto and Day (2007) (see also Ogbunugafor and Robinson (2016)). Here circles are variables that can be thought of as “stocks” (e.g., the total amount of resources (R) and consumers (C)). Symbols within circles are variables, and symbols labeling the arrows are parameters. Arrows moving towards a circle are “inflows” that increase the amount of stock, while arrows moving away from a circle are “outflows” that decrease the amount of stock. Arrows moving from one type of stock to another show how those variables are interdependent and where one type of stock is converted into another (e.g., consumed resources converted into consumer growth). In this model, arrows between two stocks represent interspecific effects, arrows that feedback on themselves represent intraspecific effects, and arrows that point out of one stock without connecting to another stock represent flows out of the system (i.e., mC).

Time series [dynamic] plot (Box Fig. 1b,c): This type of figure shows how variables of interest change over time. They can be used to assess equilibria; for example, for a given combination of parameter values, one can see how long a system exhibits transient dynamics (i.e., an initial phase of instability) before it stabilizes. They can also be used to determine if dynamics are static (i.e., population size remains constant, in Box Fig. 1b) or oscillatory (Box Fig. 1c). Finally, they

can be used to quickly assess ecological outcomes (e.g., which species persist over the long term) or how the dynamics of two variables are interrelated (i.e., synchronous vs. asynchronous dynamics).

Phase-plane [phase portrait] diagram (Box Fig. 1d,e): These summarize the trajectories of variables in multi-variable models and serve as an alternative to time series plots. In these plots, the location of the tail of each arrow marks the value of each variable at the start of a timestep, and the location of the head of the arrow is the value at the end of the timestep. The length of the arrow represents the size of the change, and the arrow's angle shows the relative magnitudes of change of one variable vs. the other. The isoclines (colored lines) represent the range of values at which the specified variable experiences no net growth if the value of the other variable were to be held constant (often not biologically plausible). The point at which the isoclines of the two species intersect is the two-species equilibrium. Here the two-species equilibrium is stable when $e=0.65$, as population sizes converge on these points. Note that variables' trajectories could instead be shown as a "vector field" in which arrows point toward stable equilibria and away from unstable equilibria.

Bifurcation plot (Box Fig. 1f): These are useful for visualizing where the nature of the dynamics changes, for example, between having a single stable value to having unstable values or to something more variable. A 'bifurcation point' is the value of a given parameter where such changes in the model's dynamical behavior occurs. In Box Fig. 1f, for example, a conversion efficiency of 0.65 leads to stable dynamics (a single consumer density over time) whereas a conversion efficiency of 0.75 causes oscillations (a range of consumer densities over time; minimum and maximum values of the oscillation are denoted on the plot). For this model, the range of abundances after the bifurcation is shaped like a tuning fork because as conversion

efficiency increases, the amplitude (or range of values covered) of the oscillating dynamics increases.

Summary plots (Box Fig. 1g): These show how changing multiple parameters at once affects an outcome of interest, for example, the type of population dynamics observed (e.g., stable limit cycles) and show the parameter values where transitions between outcomes occur. Here each axis is a different parameter in the model and the shading indicates which outcome is observed. They can also show whether or not the effect of one parameter on an outcome of interest depends on the value of other parameters. Parameters that are not explicitly shown on the axes are assigned fixed values that are often specified in figure captions (see Box Fig. 1 caption).

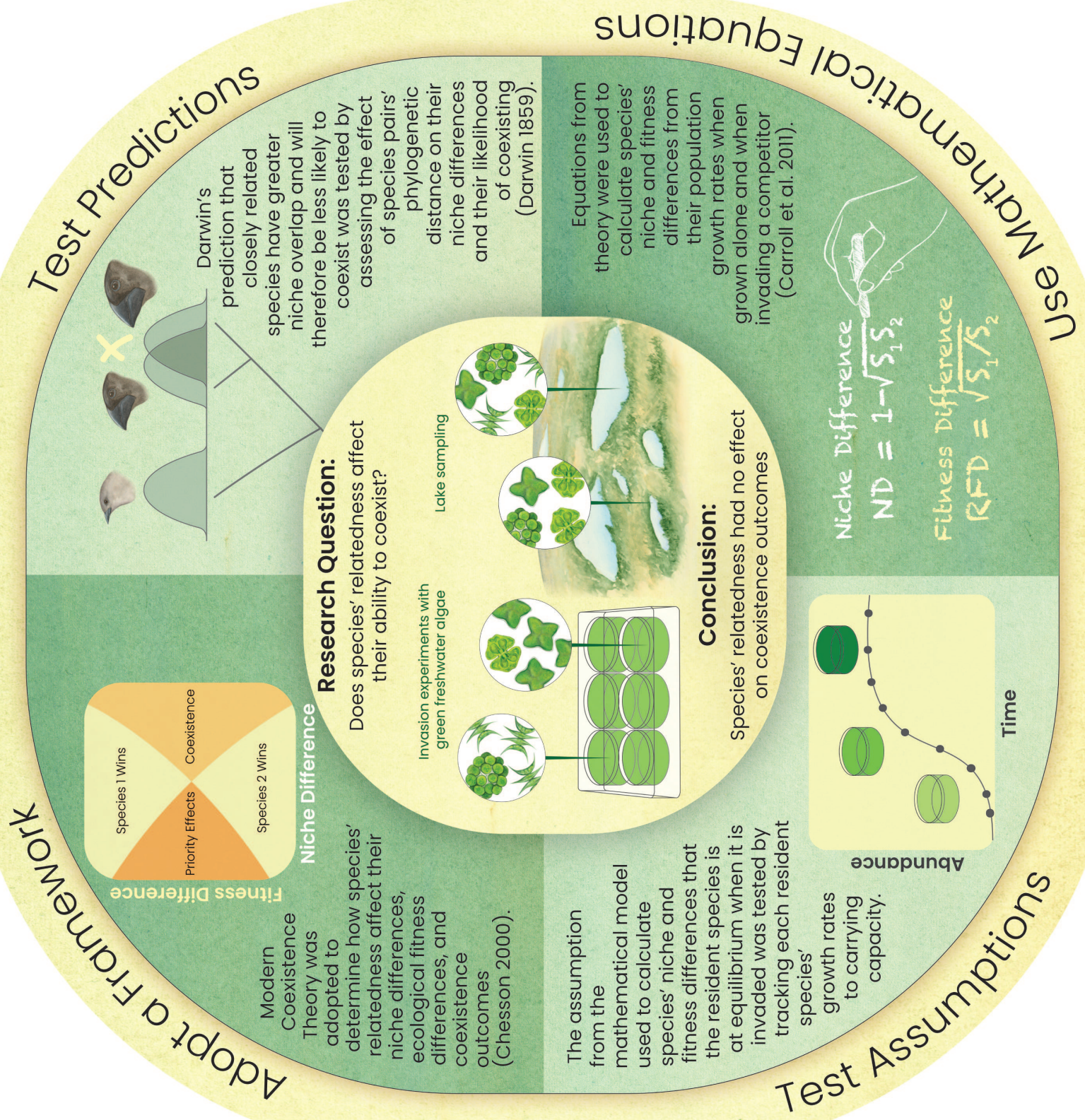
Figure legends

Figure 1. How to use theory in empirical research: a case study. The highlighted study, Narwani et al. (2013), uses theory in each of the four of the ways discussed in this paper (“How to use theory in empirical research” section). Their research question, empirical methods and conclusions are shown in the center, and the specifics of how they applied each of the four approaches are in the surrounding four sections. While Narwani et al. (2013) used all four approaches in their study, each of these four approaches can also be used effectively on its own.

Figure 2. Common theoretical notations

Box Figure 1. Five types of figures that are commonly encountered in consumer-resource theory papers: a stock and flow diagram (a), time series plots (b-c), phase plane diagrams (d-e), a bifurcation plot (f), and a summary plot (g). All panels are based on a modified Rosenzweig-MacArthur model of consumer-resource interactions (eqn. 1). Panels b-f are parameterized as follows: $r = 1$, $K = 3.5$, $a = 1.3$, $h = 0.8$, $m = 0.5$. Two values of conversion efficiency (e) are contrasted in panels b-e and are represented by dots in panels f and g. Parameters and code to reproduce this figure are available at <https://mabarbour.github.io/foodweb-theory/rosenzweig-macarthur.html>.

Figure 1



Adopt a Framework



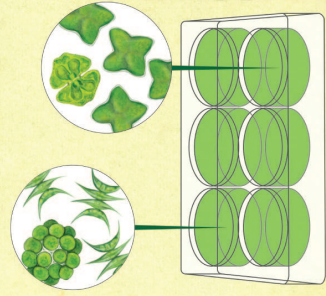
Niche Difference

Modern Coexistence Theory was adopted to determine how species' relatedness affect their niche differences, ecological fitness differences, and coexistence outcomes (Chesson 2000).

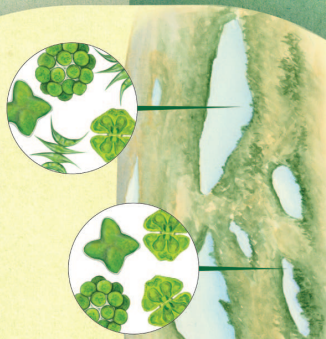
Research Question:

Does species' relatedness affect their ability to coexist?

Invasion experiments with green freshwater algae

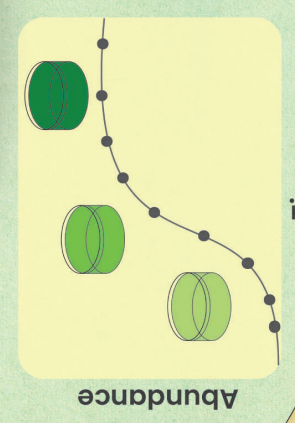


Lake sampling

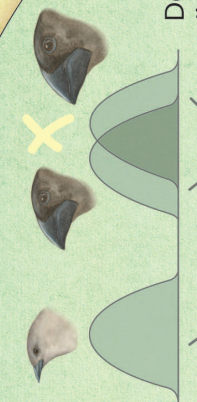


Conclusion:

Species' relatedness had no effect on coexistence outcomes



Test Predictions



Darwin's prediction that closely related species have greater niche overlap and will therefore be less likely to coexist was tested by assessing the effect of species pairs' phylogenetic distance on their niche differences and their likelihood of coexisting (Darwin 1859).

Use Mathematical Equations

Equations from theory were used to calculate species' niche and fitness differences from their population growth rates when grown alone and when invading a competitor (Carroll et al. 2011).

Niche Difference $ND = 1 - \sqrt{S_1 S_2}$

Fitness Difference $RFD = \sqrt{S_1/S_2}$

Test Assumptions

The assumption from the mathematical model used to calculate species' niche and fitness differences that the resident species is at equilibrium when it is invaded was tested by tracking each resident species' growth rates to carrying capacity.

Figure 2 (part a)

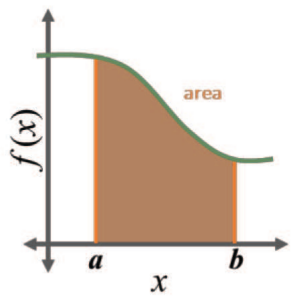
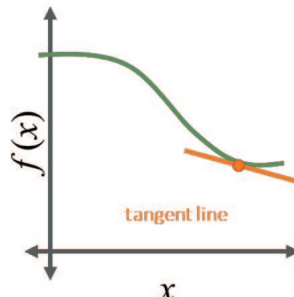
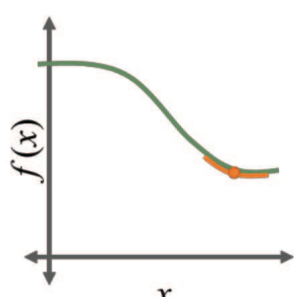
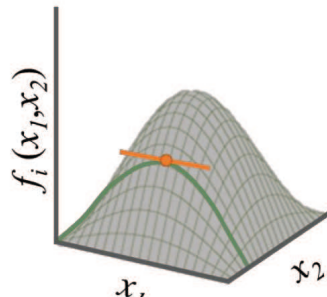
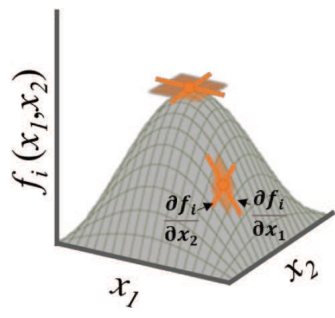
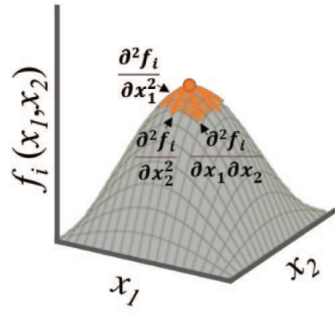
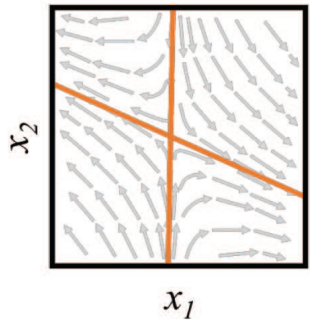
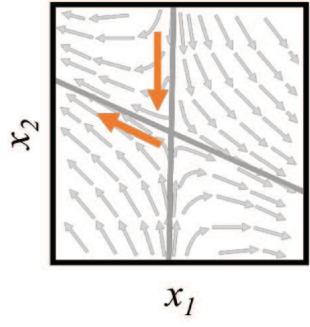
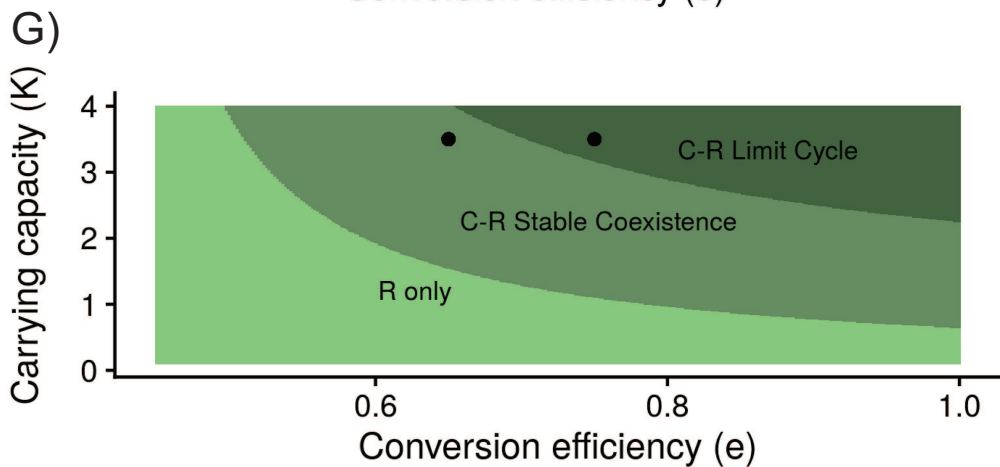
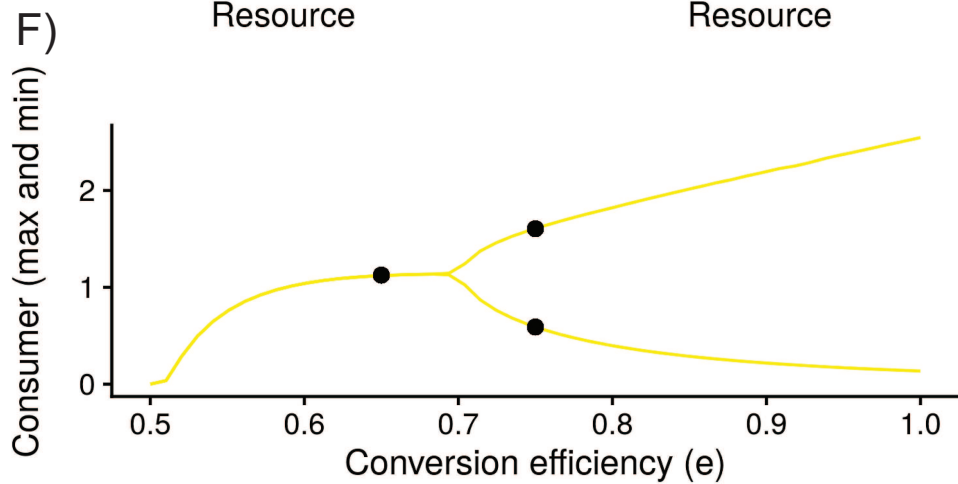
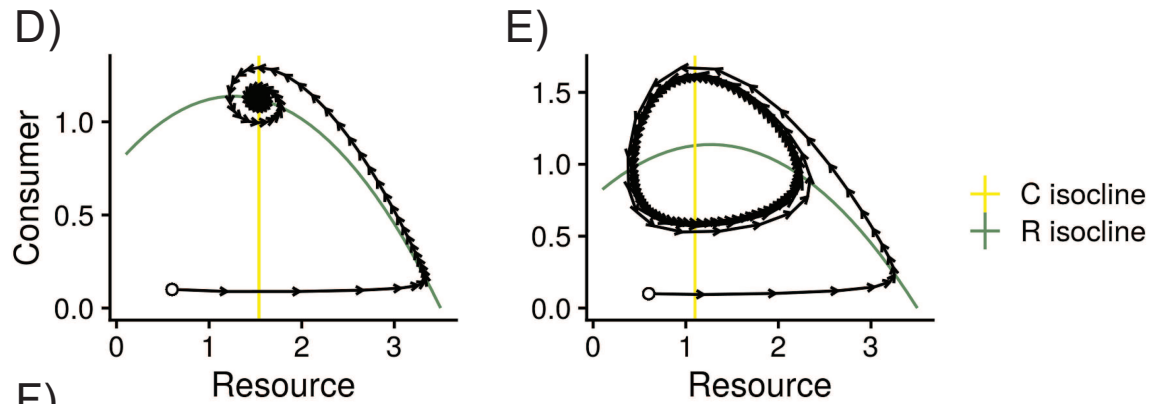
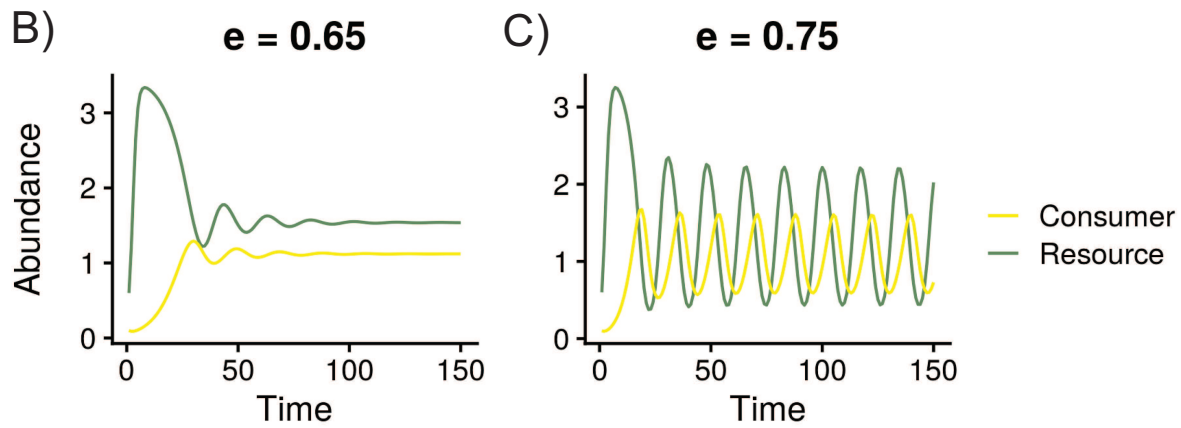
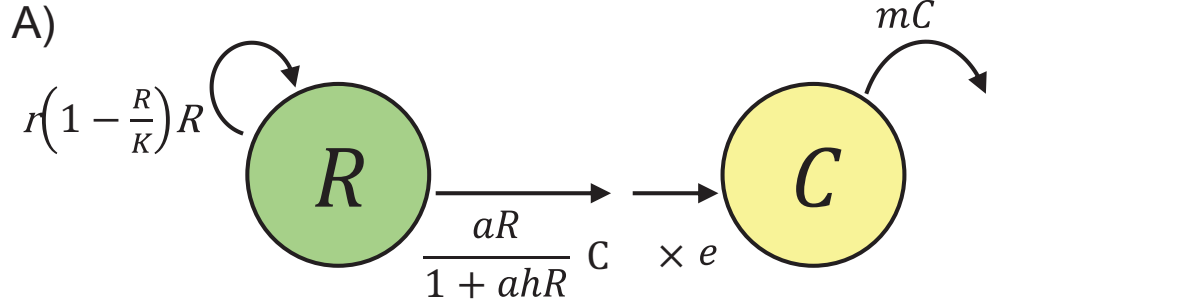
| Typical notation | Context/description | Operator | Realm | Mathematical definition | Example of use in ecology | Visual representation |
|--------------------------------|------------------------------------|--|----------------|--|--|--|
| \int | $\int_a^b f(x)dx$ | Integral | Calculus | The area under a function $f(x)$ for the range of x values between a and b . | Total amount of resources consumed by consumer, with a gradient of resource types (e.g., seeds of different sizes) (Roughgarden 1971). |  |
| d | $\frac{df}{dx}$ or $f'(x)$ | Derivative | Calculus | The slope of a function evaluated for any value of x . Equals the rise/fall in the function for a very small change in x . | Used to address rates of change of population size over time and the stability of equilibria. |  |
| d^2 | $\frac{d^2f}{dx^2}$ or $f''(x)$ | 2^{nd} -degree derivative | Calculus | The derivative of a derivative indicating how fast the slope changes (i.e., acceleration) for any value of x . | Used to describe whether fitness is stabilizing ($\frac{d^2W}{dx^2} < 0$) or disruptive ($\frac{d^2W}{dx^2} > 0$) for trait x . |  |
| ∂ | $\frac{\partial f}{\partial x}$ | Partial derivative | Calculus | When a function has more than one variable, say x and y , a partial derivative assesses the slope in one direction, say x at a specific (underived) value of y . | Change in population size over time at a particular point in space (Okubo and Levin 2013) |  |
| A or A_{ij} | Uppercase, bold Roman symbol | A matrix. If unbolded and italicized (i.e., A_{ij}), refers to a single cell in the matrix. | Matrix algebra | An i (row) by j (column) matrix. For square matrices, i and j vary from 1 to n , the number of dimensions. | A Leslie matrix contains information on fecundity and survival of different life stages. | $\mathbf{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix}$ |
| v or v_i or \vec{v} | Lowercase, Roman symbol | A list or vector if bold, a single element if unbolded | Matrix algebra | A vector | A vector representing population sizes of different life stages or species | $\mathbf{v} = [v_1, v_2, v_3, v_4]$ |

Figure 2 (part b)

| Mathematical term | Realm | Mathematical definition | Example of use in ecology | Visual representation |
|-------------------|-----------------|---|---|--|
| Jacobian matrix | Matrix calculus | A square matrix of derivatives describing linear approximations of function $f_i(x_1, x_2, \dots, x_n)$ for each equation i , with $\frac{df_i}{dx_j}$ as the entry in row i column j . | The Jacobian matrix describes the rate of change of each function with respect to each variable. Used in linear stability analyses, where the functions describe how each variable changes over time. |  A 3D surface plot of a function $f_i(x_1, x_2)$. The vertical axis is labeled $f_i(x_1, x_2)$, the horizontal axes are x_1 and x_2 . The surface is a smooth, rounded peak. Two small orange arrows on the surface represent partial derivatives: $\frac{\partial f_i}{\partial x_1}$ pointing along the x_1 axis and $\frac{\partial f_i}{\partial x_2}$ pointing along the x_2 axis. |
| Hessian matrix | Matrix calculus | A square matrix of second-order derivatives describing the curvature of function $f_i(x_1, x_2, \dots, x_n)$, with $\frac{\partial^2 f_i}{\partial x_j \partial x_k}$ as the entries in row j and column k . | Describes the local curvature of a function with many variables. Can be helpful in identifying critical points of a function (e.g., local maxima). Used in evolution to determine evolutionarily stable strategies (ESS). |  A 3D surface plot of a function $f_i(x_1, x_2)$. The vertical axis is labeled $f_i(x_1, x_2)$, the horizontal axes are x_1 and x_2 . The surface is a smooth, rounded peak. Three small orange arrows on the surface represent second-order partial derivatives: $\frac{\partial^2 f_i}{\partial x_1^2}$, $\frac{\partial^2 f_i}{\partial x_2^2}$, and $\frac{\partial^2 f_i}{\partial x_1 \partial x_2}$. |
| Eigenvector | Linear algebra | A vector associated with a square matrix which, when multiplied with the matrix, does not change its direction. | The dominant eigenvector (with the largest eigenvalue) of the Leslie Matrix is the stable age distribution of a population. |  A phase plane plot with horizontal axis x_1 and vertical axis x_2 . A vector field of small grey arrows shows the direction of change. A prominent orange line passes through the origin, representing the stable age distribution eigenvector. The arrows generally point towards this line. |
| Eigenvalue | Linear algebra | A unique real number associated with an eigenvector which represents how much the eigenvector gets scaled when multiplied by the square matrix. | Describes the direction and speed of change. The largest eigenvalue of the Leslie Matrix represents the long-term growth rate of the population. |  A phase plane plot with horizontal axis x_1 and vertical axis x_2 . A vector field of small grey arrows shows the direction and speed of change. Two orange arrows are highlighted: one pointing downwards and one pointing to the left, indicating the direction and speed of change at specific points. |



An empiricist's guide to theory

Supporting information for: An empiricist's guide to using ecological theory

The American Naturalist

Tess N. Grainger^{1,2}, Athmanathan Senthilnathan^{3*}, Po-Ju Ke^{1,4*}, Matthew A. Barbour⁵, Natalie T. Jones⁶, John P. DeLong⁷, Sarah P. Otto², Mary I. O'Connor², Kyle E. Coblenz⁷, Nikunj Goel⁸, Jawad Sakarchi², Megan C. Szojka^{2,9} Jonathan M. Levine¹ and Rachel M. Germain²

¹Princeton University, ²University of British Columbia, ³University of Tennessee, ⁴National Taiwan University, ⁵University of Zurich, ⁵University of Queensland, ⁷University of Nebraska-Lincoln, ⁸University of Texas at Austin, ⁹University of Wyoming.

*These authors contributed equally

Corresponding authors: tessgrainger@gmail.com, rgermain@zoology.ubc.ca

Table S1. List of textbook sections that introduce modeling and mathematical topics. Bolding indicates where the topic is discussed in the main text.

| Topic | Books | Sections | Notes |
|--|---------------------------|-------------------------|--|
| Graphically solving equilibria and determining their stability | Gotelli (2008) | Chapters 5, 6, Appendix | Isocline analysis for Lotka-Volterra models. |
| | Case (2000) | Chapters 5, 12, 14 | Graphical method to find equilibria. Cobwebbing and isocline analysis. |
| Analytically solving equilibria and determining their stability (Tip 5) | Case (2000) | Appendix 3 | Includes some linear algebra facts. |
| | Otto and Day (2007) | Chapters 5, 7, 8 | Detailed description of linear stability analysis for single and multiple variable models. |
| | Pastor (2008) | Chapter 2, Appendix | Basics of linear stability analysis and corresponding MatLab commands. |
| | McCann (2011) | Chapter 2 | Walkthrough of linear stability analysis. |
| Performing numerical simulations (Tip 5) | Otto and Day (2007) | Chapter 4 | Numerical simulations and plots to gain understanding of model dynamics. |
| | Crawley (2012) | Chapter 28 | Simulating discrete space/time models in R. Covers both deterministic and stochastic models. |
| Fitting models to data (“Use the mathematical equations” section) | Hilborn and Mangel (1997) | Chapters 5, 7, 9, 11 | Describes regression, likelihood, and Bayesian methods. Includes pseudocodes to do some of the analyses. |
| | Bolker (2008) | Chapters 6, 7 | Introduction to maximum likelihood |

| | | | |
|---|---------------------|---------------------|--|
| | | | and optimization methods. |
| | Roughgarden (1998) | Sections 4.4, 6.1.1 | Uses MATLAB. Gives a computational perspective. |
| | McElreath (2020) | Chapters 2, 16 | Conceptual mechanics of Bayesian data analysis. Examples of how to fit systems of ODEs to data. |
| <hr/> | | | |
| Model development | | | |
| <i>Biology to math</i> (Tips 1, 3) | Otto and Day (2007) | Chapter 2 | Detailed description on how to write out equations based on flow diagrams. |
| | Gotelli (2008) | Appendix | General description of creating continuous time population dynamics models and key steps in the analysis of such models. |
| <i>Math to biology</i> (Tips 2, 4, 5) | Otto and Day (2007) | Primer 1 | Detailed description of basic functional forms and control parameters. |
| | Bolker (2008) | Chapter 3 | Detailed description of basic functional forms and control parameters. |
| | Case (2000) | Appendix 1 Part 2 | Explains some important jargon in population models like equilibrium, variables, parameters, and initial conditions. |
| <hr/> | | | |
| General math techniques (Tips 4, 5, 6, Table 1) | Otto and Day (2007) | Appendix 1 | Commonly used algebraic rules. |
| | Pastor (2008) | Chapter 2 | Basic mathematical toolbox for commonly used techniques. |

| | | | |
|-----------------------|-----------------------------|-------------------------------|--|
| | Bolker (2008) | Appendix | Algebra and calculus basics. |
| | Bodine <i>et al.</i> (2014) | Chapter 4 | Exponential and logarithmic functions. |
| <i>Calculus</i> | Case (2000) | Appendix 1 Part 1, Appendix 5 | Visualizing mathematical functions. Properties of some typical mathematical functions with figures. Some common derivatives and integrals. |
| | Otto and Day (2007) | Appendix 2 | Basics on calculus |
| | Bodine <i>et al.</i> (2014) | Chapters 15-24 | Basic calculus: Limits, continuity, differentiation, and integration. |
| <i>Linear algebra</i> | Case (2000) | Chapter 3, Appendix 2 | Age/stage structured models (Leslie matrix). |
| | Caswell (2006) | Appendix A | Basics of matrix algebra. |
| | Otto and Day (2007) | Primer 2 | Basics of linear algebra. |
| | Bodine <i>et al.</i> (2014) | Chapters 6-9 | Matrix models; Finding eigenvalues, eigenvectors, and interpretation |
| <i>Probability</i> | Case (2000) | Appendix 6 | Describes mean, variance and standard deviation with some common probability distributions. |
| | Otto and Day (2007) | Primer 3 | Basics of probability theory. |
| | Bolker (2008) | Chapter 4 | Probability theory and distribution. |
| | Bodine <i>et al.</i> (2014) | Chapters 10-12, 25 | Basics of discrete and continuous probability spaces. Permutations and combinations. |

References

- Bodine, E.N., Lenhart, S. & Gross, L.J. (2014). *Mathematics for the life sciences*. Princeton University Press.
- Bolker, B.M. (2008). *Ecological models and data in R*. Princeton University Press.
- Case, T.J. (2000). *Illustrated guide to theoretical ecology*. Oxford University Press.
- Caswell, H. (2006). *Matrix population models: construction, analysis, and interpretation*. Second edn. Sinauer Associates.
- Crawley, M.J. (2012). *The R book*. Second edn. John Wiley & Sons.
- Gotelli, N.J. (2008). *A primer of ecology*. Fourth edn. Sinauer Associates Sunderland, Massachusetts, USA.
- Hilborn, R. & Mangel, M. (1997). *The ecological detective: confronting models with data*. Princeton University Press.
- McCann, K.S. (2011). *Food webs (MPB-50)*. Princeton University Press.
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press.
- Otto, S.P. & Day, T. (2007). *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press.
- Pastor, J. (2008). *Mathematical ecology of populations and ecosystems*. John Wiley & Sons.
- Roughgarden, J. (1998). *Primer of ecological theory*. Pearson.