Disturbance, productivity, and species diversity: empiricism vs. logic in ecological theory

MICHAEL A. HUSTON

Department of Biology, Texas State University, San Marcos, Texas 78666 USA

Abstract. The “intermediate disturbance hypothesis” and the “intermediate productivity hypothesis” have been widely recognized concepts for explaining patterns of species diversity for the past 40 years. While these hypotheses have generated numerous reviews and meta-analyses, as well as persistent criticism, two prominent papers have recently concluded that both of these hypotheses should be abandoned because of theoretical weaknesses and failure to predict observed diversity patterns. I review these criticisms in the context of the continuing tension between logic and empiricism in the development of ecological theory, and conclude that most of the criticisms are misguided because they fail to recognize the inherent connections between these two hypotheses, and consequently fail to test them appropriately. The logic of every hypothesis is based on the underlying assumptions. In the case of these two hypotheses, the assumptions on which the criticisms of their logic depend are falsified by the strong empirical support for the linked predictions of the hypotheses. This conclusion calls for a reevaluation of the basic assumptions upon which most of ecological competition and diversity theory is based.

Key words: disturbance; diversity; dynamic equilibrium model (DEM); equilibrium; experiments; gradients; intermediate disturbance hypothesis (IDH); intermediate productivity; mortality; NPP; productivity; theory.

INTRODUCTION

Recent critiques have called for the rejection of two of the best-known hypotheses to explain spatial patterns of species diversity, the Intermediate Disturbance Hypothesis, IDH (Grime 1973a, b, Horn 1975, Connell 1978, J. F. Fox 1979, Wilkinson 1999), and the Intermediate Productivity Hypothesis, IPH (Grime 1973a, b, 1979), also called the Humped-back Model. The IDH is best known from Connell’s (1978) heuristic description published with supporting data from rain forests and coral reefs, although it previously had been presented graphically by Grime (1973a) and Horn (1975). The IPH was originally proposed as an explanation for patterns of species diversity measured along a gradient of herbaceous plant communities ranging from low to high productivity (Grime 1973a, b, Al-Mufti et al. 1977).

Jeremy W. Fox (2012) provided a detailed critique of what he considered to be the failures of the intermediate disturbance hypothesis. He argued that the IDH fails both in its ability to predict empirical patterns and in the logical failure of its presumed mechanisms to produce stable coexistence under the “intermediate” disturbance conditions where species diversity is predicted to be highest. Adler et al. (2011) evaluated the IPH with data from a global sampling network and concluded that “productivity is a poor predictor of plant species richness,” on the basis of the empirical failure of their data to match the predicted patterns. Others have criticized the theoretical inadequacy of both productivity and disturbance as explanations for long-term stable equilibrium of high species diversity (Abrams 1995, Chesson and Huntly 1997, Chesson 2000a, b, Shea et al. 2004).

These recent challenges to well-established ecological hypotheses highlight a fundamental disagreement about how ecological processes operate and how ecological theories should be developed and tested. This disagreement is between the use of empiricism vs. logic as the primary criterion for the development and evaluation of ecological theory. Empiricism is based on data and leads to hypotheses intended to explain and predict the actual patterns observed in nature. Logic focuses on the...
mechanistic consistency of theories in relation to assumptions about how nature operates and leads to hypotheses about what patterns are theoretically possible in nature. Empiricism and logic have led to two very different worldviews about how natural systems operate, which are illustrated by the arguments about the effects of productivity and disturbance on species diversity, as well as related arguments about the effects of species diversity on productivity (Kareiva 1994, 1996, Naæem et al. 1994, 1996, Huston 1997, Huston et al. 2000, Wardle et al. 2000).

Fox’s summary of specific mechanisms and mathematical criteria that logically can or cannot contribute to stable long-term coexistence of species is clear and informative. However, he ignores the fact that examples of many of these mechanisms were discussed by Connell (1978) in his description of the IDH, even though the IDH emphasized the short-term prevention of competitive exclusion, rather than stable long-term coexistence. Both Fox and Adler et al. make the same fundamental error that has been made in virtually all commentaries, studies, and meta-analyses of both the IDH and IPH. This error is the assumption that the effects of disturbance on diversity are independent of the effects of productivity on diversity, and vice versa, which results in both the IDH and the IPH being evaluated as single-factor models. This error invalidates most of the empirical studies done on the IDH and IPH and helps to explain the apparently poor predictive capability of both models.

Species Diversity Hypotheses: Comparisons and Issues

In proposing his version of the IDH, Connell (1978) emphasized the universality of the “intermediate” position of the diversity maximum along a gradient from low to high disturbance (mortality-causing events such as predation, wave damage, or severe windstorms). Connell also explicitly rejected Huston’s (1979) hypothesis that productivity or population growth rates could affect the response to disturbance, such that the maximum level of diversity along a disturbance gradient would shift from one end of the gradient to the other depending on how fast the populations were growing and able to recover from the disturbance (Connell 1978:1305).

Because the IDH as presented by Connell was strictly a verbal model, J. W. Fox focused his critique of the IDH on a single output of a set of computer simulations from which the IDH can be derived (Huston 1979: Figs. 3 and 5). The full set of computer simulations of the nonequilibrium dynamics of the Lotka-Volterra competition equations, from which Fox drew his example of the IDH, was originally used to illustrate the interaction of growth rates (productivity) and mortality (disturbance) in their effects on species diversity. This interaction is the key process underlying the dynamic equilibrium model or DEM (Huston 1979, 1994).

In the context of the DEM, the IDH is a special case (Fig. 1Db) that occurs only under certain conditions, specifically where productivity and potential population growth rates are also intermediate. If growth rates are very low, intermediate or high levels of mortality can drive populations to extinction, and the highest diversity is usually found at low levels of mortality (Fig. 1A). However, if growth rates are high, the maximum level of diversity is typically found where disturbance frequencies and/or intensities are also high because competitive exclusion occurs rapidly and high rates of mortality are required to prevent competitive exclusion (Fig. 1B).

Thus, the effect of mortality-causing disturbances on diversity reverses between low-productivity environments and high-productivity environments.

While the publication of the DEM less than a year after Connell’s IDH made it clear that the IDH was a special case that could only occur under specific conditions, most experimental and observational studies of the IDH have ignored the effects of productivity (e.g., Feminella and Hawkins 1995, Mackey and Currie 2001). Consequently, some disturbance studies have concluded that there was not only no intermediate disturbance effect, but also that disturbance had no predictable effect on species diversity at all (Feminella and Hawkins 1995, Steinman 1996). It is only when disturbance effects on diversity are stratified by productivity that the consistent effects of mortality on species diversity can be observed, including some cases of an “intermediate disturbance” response (Huston 1979, 1994, 2002, Proulx and Mazumder 1998, Kershaw and Mallik 2013). Numerous meta-analyses and reviews of the effect of disturbance on species diversity have concluded that nonsignificant, linear increasing, and linear decreasing responses of diversity to disturbance are more frequent than the unimodal IDH pattern (e.g., Feminella and Hawkins 2000, 2001). This is not surprising, given the prediction of the DEM that all three types of patterns should be found (Fig.1Da–c), depending on the productivity of the systems where the effects of mortality-causing disturbances were measured.

Thus, the supposed empirical failure of the IDH is not due to the failure of its logic, but rather because the failure to test the IDH under known productivity conditions meant that any intermediate diversity maxima that may have been present were obscured by other diversity maxima at different levels of disturbance in data sets that ignored variation in productivity (Huston 2002); see Fig. 1Dg. The pattern predicted by the IDH does occur, but it can only be detected within a specific range of productivity, when the data are analyzed using appropriate stratification by productivity and other factors such as species pool size.

Of the studies cited by Fox as empirical falsifications of the IDH, all were experimental manipulations of microorganisms conducted under controlled conditions of food availability and disturbance dynamics (Warren 1996, Buckling et al. 2000, Scholes et al. 2005,
FIG. 1. Contrasting effects of disturbance on species diversity, depending on productivity (population growth rates) of the system. (A) Effect of mortality-causing disturbances on plant species richness in nutrient-poor ecosystems (aquatic and marine) in which mortality is caused by grazing (Proulx and Mazumder 1998). These conditions correspond to graph (a) in (D). All of the studies showed a significant decrease in species richness under unproductive conditions, as predicted by the dynamic equilibrium model, DEM. (B) Effect of mortality caused by grazing or mowing on plant species richness in nutrient-rich ecosystems (natural or fertilized, aquatic, marine, and terrestrial) (Proulx and Mazumder 1998). These conditions correspond to graph (c) in (D). Results are shown for 24 comparisons reported in 13 published studies examined in the literature review by Proulx and Mazumder (1998). Most (21 of 24) showed an increase in species richness under productive (i.e., nutrient-rich) conditions as predicted by the DEM. Plus signs indicate a significant positive response (14 out of 24 comparisons), with a significant negative response (−) in one out of 24 comparisons. Numbers by lines refer to study numbers in Proulx and Mazumder (1998). Note that study 22 and, to a lesser extent, studies 21a and b, show the unimodal response of the intermediate disturbance hypothesis, corresponding to graph (b) in (D). (C) Predictions of the dynamic equilibrium model, DEM (Huston 1979, 1994), showing the effects on species diversity of the
NPP is incorrect at high levels of precipitation and AET, rather than positively correlated as the authors assumed. Correlated with NPP at high levels of the surrogate, AET (actual evapotranspiration), that are negatively correlated with surrogates of NPP, typically precipitation or measurements of net primary productivity (NPP), but including tropical forests are not based on actual measurement periods, and so did not clearly falsify the IDH. 

In contrast to the IDH, the intermediate productivity hypothesis (IPH), also called the humped-back model (Grime 1973a, b, 1979), has been consistently supported by empirical studies (Grime 1973a, Al-Mufli et al. 1977, Silvertown 1980, Moore and Keddy 1989, Wheeler and Shaw 1991, Gough et al. 1994, Kull and Aan 1997, Guo and Berry 1998, Mittelbach et al. 2001, Jacquemyn et al. 2003, Virtanen et al. 2012). The original studies that led to the IPH were conducted in British pastures and herbaceous plant communities in which productivity could be easily measured by cutting and weighing the plants at the end of the growing season (e.g., Al-Mufli et al. 1977). These communities have a relatively uniform disturbance regime related to annual cycles of mowing and grazing. Across a broad gradient of productivity (150–2600 g living and dead dry plant mass/m²), the data revealed a clear unimodal (quadratic) pattern (Fig. 2A). However, contradicting the generality of the unimodal IPH pattern, several recent diversity–productivity reviews have identified an increase in the frequency of monotonically increasing productivity–diversity patterns at larger scales (regional and continental to global, Mittelbach et al. 2001, Wright et al. 2006, 2010), particularly with use of climate surrogates for NPP (Mittelbach et al. 2001, Gillman and Wright 2006). Most of these results are flawed and irrelevant to the general validity of the IPH and DEM. 

Most analyses that find increasing plant diversity with “increasing productivity” at large spatial scales (usually including tropical forests) are not based on actual measurements of net primary productivity (NPP), but rather on surrogates of NPP, typically precipitation or AET (actual evapotranspiration), that are negatively correlated with NPP at high levels of the surrogate, rather than positively correlated as the authors assumed (Huston 2012). Consequently, the assumption that precipitation and AET are positively correlated with NPP is incorrect at high levels of precipitation and AET, where productivity is incorrectly presumed to be highest. This results from the effects of warm, wet conditions on the weathering and leaching of all major plant nutrients from the soil, which inevitably reduces soil fertility and plant productivity unless geological processes provide new material with high nutrient content for soil formation (Jenny and Leonard 1934, Albrecht 1940, 1957, Sanchez 1976, Walker and Syres 1976, Huston 2012). Actual measurements of the NPP of tropical rain forests (Jordan 1971a, b, Willson 1973, Cramer et al. 2001, Schuur 2003, Huston and Wolverton 2009, Huston 2012, Clark et al. 2013) demonstrate that they have relatively low productivity, and thus their diversity patterns are consistent with the predictions of the IPH and DEM that plant diversity is usually highest at relatively low levels of productivity. 

Well-designed studies that measure plant diversity across a broad range of productivity and control for disturbance dynamics typically find a clear unimodal response (although the mode may be shifted to higher or lower productivity depending on the disturbance regime). In addition to challenges related to the measurement of productivity in different life-forms of plants (Clark et al. 2001), at least three other factors may obscure the relationship between species diversity and productivity or disturbance.

The first major problem is that, for any disturbance gradient, there is expected to be a unimodal diversity pattern, with maximum diversity at intermediate levels of disturbance only under intermediate conditions of productivity (Fig. 1Db). Unless the effects of disturbance on diversity are sampled under known conditions of productivity and the analysis of disturbance effects is stratified by productivity level, fragments of the full pattern could potentially appear as monotonically increasing (Fig. 1Dc), monotonically decreasing (Fig. 1Da), or unchanging if measured near the level of maximum diversity (Fig. 1Db at intermediate mortality rate). Mixing data on the effect of disturbance on diversity without stratifying by productivity may result in superimposing contrasting patterns, which produces a scatter of data with no clear pattern (Fig. 1Dg), leading to the incorrect conclusion that disturbance does not have a predictable effect on species diversity. 

A second, related problem is that the effect of productivity on diversity is highly dependent on the disturbance–mortality regime. The effect of productivity may be an increase in diversity with increasing...
FIG. 2. Patterns of plant species richness along gradients of plant productivity, showing the typical pattern of highest plant diversity near the low end of the productivity gradient. Productivity is either estimated from direct measurements of plant biomass or approximated using soil fertility as a surrogate for plant productivity. These studies did not explicitly measure disturbance regime, but by sampling a specific stage of plant succession, e.g., pastures and grasslands vs. closed-canopy forests, they constrain all sites within a relatively narrow range of disturbance frequencies. (A) Herbaceous plant species richness in relation to live and dead plant biomass in British pastures and grasslands (Al-Mufti et al. 1977). (B) Predicted differences in the pattern of species diversity along a productivity gradient caused by differences in the number of species available in regional species pools (Huston 1999). (C) Combined data set of species numbers in relation to natural variation in live plant biomass for 48 sites, with 36 in North America, and 12 scattered across Europe, Australia, Africa, and China (Adler et al. 2011). Species pool sizes among the sites ranged from about 5 to 45 species. All sites combined (gray dots plus black dots) have no statistically significant linear or nonlinear relationships between biomass and species richness. However, the black dots, which represent maximum species richness in relation to biomass (estimated as the 20 plots with the highest species richness in each 100 g/m² segment) show a strong quadratic relationship that is qualitatively similar to most other studies of diversity along productivity gradients (e.g., panels A, D, E, and F).
productivity, or a decrease in diversity with increasing productivity, or a unimodal pattern, depending on the rate of mortality (Fig. 1Dd–f). If data are collected along a productivity gradient from sites that differ in disturbance history, the actual effect of productivity on diversity may be completely obscured so that no pattern can be detected (Fig. 1Dh). Also, the full response of diversity to productivity can only be detected if the diversity is measured across the entire gradient (e.g., Al-Mufti et al. 1977, Guo and Berry 1998, Huston 2002, Fridley et al. 2012) under consistent disturbance regimes. While some meta-analyses and reviews of studies that evaluated the effects of productivity on diversity in different plant communities have found that unimodal “humped” patterns occurred most frequently, responses that are positive, negative, or nonsignificant are also found (Wright et al. 1993, Mittelbach et al. 2001). This variability is not surprising, given that the productivity studies were not stratified by mortality regimes, and consequently could not identify the contrasting responses predicted by the DEM to occur under different mortality conditions.

A third major problem that leads to invalid tests of these hypotheses is that the predictions of the IDH and IPH, as well as of the full DEM, are predictions of relative diversity, not absolute diversity. The DEM predicts that variation in the rate of competitive exclusion leads to variation in the probability that species can potentially coexist. Under conditions in which competitive exclusion is slowed or prevented, most of the species in the local species pool can potentially coexist, at least for the relatively short term of 10s to 100s of years, with the maximum number of co-occurring species set by the size of the local species pool (Fig. 2B). Diversity is predicted to be low under two other conditions due to two completely different mechanisms. (1) Where productivity is very low, diversity is predicted to be low due to the failure of most species to survive under conditions of very low growth rates, potentially with the added impacts of mortality-causing disturbances. (2) Where productivity is high and disturbance frequency low, diversity is also predicted to be low, but in this case it is because of high rates of competitive exclusion where plant populations are growing rapidly and competing intensely. The low number of species typically found under these two extreme conditions is unrelated to the size of the local species pool (Fig. 2B).

Thus, the maximum number of potentially competing species that could co-occur at the local spatial scale is determined by the total number of species within the region, which is called the “local” or “regional species pool,” depending on the temporal and spatial scales being considered. The size of regional species pools is determined by a variety of environmental, historical, and biogeographical contingencies, and usually differs between regions (Ricklefs 1987, Ricklefs and Schluter 1993). Consequently, studies that combine data from productivity gradients in multiple different regions produce a distribution of data that differs predictably from the patterns found in studies limited to a single region.

Intraregional productivity gradients, where all sites potentially share the same species pool, often produce a curvilinear diversity pattern that can be fit by a quadratic or higher order equation, with diversity decreasing at the highest levels of productivity (e.g., Fig. 2A, D–F). In contrast, interregional studies that combine the data from productivity gradients in multiple regions that differ in the size of the species pool produce a “filled curve” in which many of the values of species diversity at a specific level of productivity fall below a maximum level that represents the highest regional pool size across the multiple regions (Fig. 2C). When such filled curves are evaluated using traditional linear or nonlinear regressions, the statistical analysis typically reveals that there is no significant pattern.

Traditional linear regression and ANOVA methods are often inappropriate for analysis of ecological data because the data fail to meet the assumptions of the statistical model. The key issue is interpretation of the variance in the data. These statistics assume that variance in the data is the result of random errors (“noise” or unmeasured factors) that cause some of the values of the dependent variable to be higher or lower than the true relationship between the independent and dependent variables, which is considered to be the mean, e.g., the linear regression line. This assumption requires that the variance of the “error” in the dependent variable have the same distribution across the full range of the independent variable, a condition called “homoscedasticity.” Many types of ecological data are not homoscedastic, not only because the distribution of the errors differs across the range of the independent variable, but often also because the actual ecological processes that produce the variance differ over the range of values of the independent variable (Huston 2002, Huston and McBride 2002).

The study conducted by Adler et al. (2011:1750), which concluded that “productivity is not a good predictor of plant species richness” was an interregional
Table 1. Summary of published experimental tests of the dynamic equilibrium model (DEM); accepted experiments must have manipulated both productivity and disturbance, and have a minimum of three levels of each experimental factor to test for nonlinear (unimodal) responses in that factor.

<table>
<thead>
<tr>
<th>System</th>
<th>IDH uni-modal</th>
<th>IPH uni-modal</th>
<th>Productivity levels</th>
<th>Disturbance levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed prairies</td>
<td>yes</td>
<td>no</td>
<td>three topographic positions</td>
<td>three grazing intensities</td>
</tr>
<tr>
<td>Old-field plants</td>
<td>yes</td>
<td>yes</td>
<td>four levels fertilizer addition</td>
<td>four mowing frequencies</td>
</tr>
<tr>
<td>Continuous culture</td>
<td>no</td>
<td>no</td>
<td>three levels nutrients</td>
<td>three levels mortality</td>
</tr>
<tr>
<td>Marine benthic mesocosms</td>
<td>yes</td>
<td>yes</td>
<td>seven levels detrital carbon addition</td>
<td>seven frequencies of mechanical stirring</td>
</tr>
<tr>
<td>Old-field plants</td>
<td>yes</td>
<td>no</td>
<td>four fertilization levels</td>
<td>four levels of annual tilling (0–100%)</td>
</tr>
<tr>
<td>Subtidal algae</td>
<td>NA</td>
<td>yes</td>
<td>eight nutrient levels: four enrichment levels at two contrasting sites</td>
<td>two levels grazing (open vs. caged)</td>
</tr>
<tr>
<td>Simulated tree holes</td>
<td>yes, 2/12</td>
<td>NA</td>
<td>two levels detrital carbon addition</td>
<td>three frequencies of drying plus three predator densities</td>
</tr>
<tr>
<td>Experimental adaptive radiation in microbes</td>
<td>yes</td>
<td>yes</td>
<td>eight levels serial nutrient dilution 1000× range</td>
<td>five frequencies of mixing and transfer of 0.001 volume to fresh medium</td>
</tr>
<tr>
<td>Protists in laboratory</td>
<td>yes</td>
<td>yes</td>
<td>six levels wheat grains</td>
<td>five frequencies of 70% mortality</td>
</tr>
<tr>
<td>Marine hard substrate</td>
<td>yes</td>
<td>no</td>
<td>three nutrient concentrations</td>
<td>six frequencies of scraping</td>
</tr>
<tr>
<td>Subtidal algae</td>
<td>NA</td>
<td>yes (deep water only)</td>
<td>three nutrient levels</td>
<td>two levels grazing (open vs. caged), two depths (&lt;1 m, &gt;2.5 m)</td>
</tr>
<tr>
<td>Aquatic microcosms</td>
<td>NA</td>
<td>NA</td>
<td>−two levels nutrients</td>
<td>two frequencies and three intensities of mortality</td>
</tr>
<tr>
<td>Sheep-grazed pasture</td>
<td>no</td>
<td>no</td>
<td>five fertilization levels</td>
<td>three grazing levels</td>
</tr>
</tbody>
</table>

Notes: There were four criteria for evaluating consistency of experimental results with theoretical predictions. (1) DEM disturbance × productivity interaction, in which the slope of all or part of the response of diversity to a gradient of productivity (or mortality) changed between different levels of mortality (or productivity). For example, in Fig. 1D the slope of the effect of mortality on species diversity on the left side of the graph (lower mortality rates) changes between panels (a) and (b). A shift from a high negative (or positive) slope to a low negative (or positive) slope also meets this criterion. All studies included here meet this criterion. (2) DEM Reversal, which is a shift from a negative to a positive slope in all or part of the response of diversity to a gradient of productivity (or mortality) between different levels of mortality (or productivity). For example, in Fig. 1D the comparisons a–b, b–c, a–c, d–e, e–f, and d–e meet this criterion. Note that if a study meets criterion 2 it also meets criterion 1, but not vice versa. The first eight studies included here meet this criterion, but the last four do not. (3) IDH (intermediate disturbance hypothesis) uni-modal: “yes” means there is a uni-modal response of diversity along a mortality gradient similar to Fig. 1Db. The maximum diversity (mode) may be shifted toward either higher or lower rates of mortality and is not limited to a central position. Study 22 in Fig. 1B meets this criterion. (4) IPH (intermediate productivity hypothesis) unimodal: “yes” means there is a unimodal response of diversity along a productivity gradient similar to Fig. 1De. The maximum diversity (mode) may be shifted toward either higher or lower rates of productivity and is not limited to a central position. Data illustrated in Fig. 2 A, C, D, E, and F all meet this criterion.

project that combined data from natural productivity variation within 48 sites distributed across five continents (Fig. 2C). One reason that productivity did not seem to be a good predictor of species diversity was that the data were from multiple sites that had species pools of a wide range of sizes (approximately 5–45), which were combined and analyzed inappropriately, in violation of the assumptions of the model they claimed to be testing (see Fridley et al. 2012). A second reason was the failure to consider the effects of disturbance, as a result of measuring only live biomass and ignoring standing dead biomass, which can have a strong “competitive” effect in grasslands through suppressing seedling germination and the regrowth of perennial plants (Carson and Peterson 1990, Facelli and Pickett 1991). A third reason was the failure to sample uniformly across the full range of productivity in the study. Most of their samples were from a relatively small portion of the productivity gradient, at low levels of productivity where the highest diversity is typically found. This clustered sampling produced a bivariate distribution of diversity and productivity in which the shape of the maximum diversity curve could be misinterpreted as a statistical artifact (e.g., Grace et al. 2012). Furthermore, the linear and nonlinear statistical analyses of Adler et al. completely ignored the extreme heteroscedasticity of their data, which was caused by all data values falling on or below a unimodal upper boundary, described by a
quadratic equation (gray line in Fig. 2C). Standard statistical methods are completely inappropriate for data of this type, where alternative methods, such as quantile regression or upper boundary analysis may provide a more informative description of the true relationship between the dependent and independent variables.

As in the critique by Fox (2012), Adler et al. (2011) also noted that most of the theoretical models for stable long-term coexistence of species predict that diversity should increase monotonically with increasing productivity, rather than producing a unimodal pattern of species diversity along productivity gradients (e.g., Abrams 1995). The fact that consistent unimodal and decreasing patterns of diversity in relation to productivity are widely reported in spite of being inconsistent with the predictions and coexistence criteria of equilibrium competition theory brings us to the critical question of how ecological theory should be developed and evaluated.

THE EVALUATION OF DIVERSITY THEORY: LOGIC VS. EMPIRICISM

Fox (2012) argues that the “disturbance slows or prevents competitive exclusion” mechanism of the IDH fails to produce long-term stable coexistence as defined by equilibrium analysis of competition equations (e.g., Abrams 1983, 1995, Chesson and Huntly 1997, Chesson 2000b). The theoretical “problem” is that, logically and as demonstrated by mathematical analysis, neither simple density-independent disturbances, nor low productivity, nor environmental fluctuations can by themselves prevent competitive exclusion from occurring under stable conditions over long time periods (Chesson and Huntly 1993, 1997, Abrams 1995, Chesson 2000b, Fox 2012). Fox asserts that this logical failure invalidates the IDH sufficiently that it should be abandoned, and he endorses the similar conclusions reached by Adler et al. (2011) about the IPH. He also claims that this same logical failure applies to the simple simulation model of the Lotka-Volterra competition equations that was used to illustrate the predictions of the DEM (Huston 1979).

However, the DEM and its predictions based on the interaction of productivity and disturbance, including the “intermediate disturbance effect” and the “intermediate productivity effect” have been extensively validated with observational and experimental data (Huston 1980b, 1994; Fig. 5.9, Rashit and Bazin 1987, Proulx and Mazumder 1998, Worm et al. 2002, Kneitel and Chase 2004, Scholes et al. 2005, Svensson et al. 2007, Haddad 2008).

Table 1. Extended.

<table>
<thead>
<tr>
<th>Duration</th>
<th>Focal trophic level</th>
<th>No. trophic levels</th>
<th>No. plant spp.</th>
<th>No. animal spp.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 yr</td>
<td>grasses</td>
<td>1</td>
<td>14+</td>
<td>1</td>
<td>Tomanek and Albertson (1957); unpublished data analysis by M. Huston analysis by M. Huston</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Huston 1980b, 1994: Fig. 5.9</td>
</tr>
<tr>
<td>3 yr</td>
<td>plants</td>
<td>1</td>
<td>~95</td>
<td>0</td>
<td>Huston 1980b, 1994: Fig. 5.9</td>
</tr>
<tr>
<td>6 weeks</td>
<td>bacteria, protozoan flagellates, protozoan predators marine benthic infauna</td>
<td>3</td>
<td>0</td>
<td>not provided</td>
<td>Rashit and Bazin (1987)</td>
</tr>
<tr>
<td>12 weeks</td>
<td>marine benthic infauna</td>
<td>?</td>
<td>1 dried brown alga for carbon source</td>
<td>81 taxa marine infauna</td>
<td>Widdicombe and Austen (2001)</td>
</tr>
<tr>
<td>7 yr</td>
<td>plants</td>
<td>1</td>
<td>15 herbaceous</td>
<td>0</td>
<td>Wilson and Tilman (2002)</td>
</tr>
<tr>
<td>11 months</td>
<td>plants</td>
<td>1</td>
<td>~5</td>
<td>~6</td>
<td>Worm et al. (2002)</td>
</tr>
<tr>
<td>23 d</td>
<td>bacteriovores</td>
<td>1</td>
<td>1 dried ash leaves for carbon source</td>
<td>16 protozoans, 3 rotifers, 1 mosquito larva as predator</td>
<td>Kneitel and Chase (2004)</td>
</tr>
<tr>
<td>16 d</td>
<td>bacteria</td>
<td>1</td>
<td>0</td>
<td>14 colony morphs of Pseudomonas fluorescens evolved from isogenic inoculum</td>
<td>Kassen et al. (2004)</td>
</tr>
<tr>
<td>24 weeks</td>
<td>bacteriovores</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>Scholes et al. (2005)</td>
</tr>
<tr>
<td>24 weeks</td>
<td>mixed</td>
<td>3+</td>
<td>15 macroalgae</td>
<td>17 (7 phyla)</td>
<td>Svensson et al. (2007)</td>
</tr>
<tr>
<td>6 months</td>
<td>plants</td>
<td>1</td>
<td>7 macroalgae</td>
<td>4 gastropods, 4 amphipods, 2 isopods</td>
<td>Korpinen et al. (2007)</td>
</tr>
<tr>
<td>31 d (30–120 generations)</td>
<td>bacteriovore</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>Haddad et al. (2008)</td>
</tr>
<tr>
<td>27 yr</td>
<td>plants</td>
<td>1</td>
<td>20 herbaceous</td>
<td>1 sheep</td>
<td>Laliberté et al. (2012)</td>
</tr>
</tbody>
</table>

September 2014

EMPIRICISM VS. LOGIC

2389

CONCEPTS & SYNTHESIS
This is particularly significant because the DEM made a priori predictions that the effects of disturbance on diversity would reverse depending on the level of productivity, and the effects of productivity on diversity would reverse depending on the disturbance regime. These phenomena had never been previously reported, and were not discovered until data were appropriately stratified and analyzed, as in the papers just cited.

Much of the criticism of the DEM, and the IDH and IPH as elements of the DEM, is a consequence of misrepresentation of these hypotheses by most of their critics. These hypotheses were not presented as precise, completely defined models, but rather as concepts illustrated either by the heuristic graphs and verbal descriptions of the IDH (Grime 1973a, Connell 1978) and the IPH (Grime 1973a) or the computer simulation model of the DEM (Huston 1979). The essential feature of the DEM simulation model was not that either disturbances or low productivity produced stable long-term coexistence (they clearly did not; see Huston 1979: Figs. 3 and 5), but rather that disturbances or low productivity delayed competitive exclusion, potentially allowing many other coexistence mechanisms to operate and maintain a higher level of diversity for a longer period under some conditions than under other conditions. Fox acknowledges this point, and also acknowledges that the temporal environmental variations described by Hutchinson (1961) would allow temporary coexistence by delaying competitive exclusion, although they could not prevent competitive exclusion under stable long-term equilibrium conditions without the operation of additional mechanisms.

There is little doubt that competitive exclusion can occur rapidly under some conditions. The key to understanding high diversity is being able to predict when competitive exclusion is slowed sufficiently to allow prolonged coexistence and the operation of other coexistence-promoting mechanisms, which may or may not be equivalent to stable long-term coexistence under the equilibrium conditions of mathematical theory.

Fox has created oversimplified caricatures of both the IDH and DEM by first ignoring all of the coexistence-promoting mechanisms actually discussed by Connell (1978) and Huston (1979), and then criticizing the original hypotheses because his caricatures did not include some obvious, as well as some more subtle, mechanisms that allow coexistence. Fox points out that neither Connell’s disturbances nor Hutchinson’s environmental conditions are sufficient to produce coexistence because “fluctuations around averages simply cancel out” over long time periods (Fox 2012: Box 1). Rather, Fox argues, “trade-offs between traits that confer high fitness when population densities are high and resources are scarce, and traits that confer high fitness when population densities are low and resources are abundant, can indeed promote competitive coexistence that would not occur in undisturbed environ-
ments” Fox 2012: Box 2). Fox adds that “… such trade-offs generate disturbance-mediated coexistence via nonlinearities and non-additivities, not because disturbances simply reduce species’ densities” (Fox 2012:1235; see Levins 1979, Chesson and Huntly 1997, Grover 1997, Chesson 2000b).

However, both Connell (1978) and Huston (1979) have extensive discussions of the importance of specific mechanisms of the types that Fox claims are missing from the hypotheses (Fox 2012: Boxes 2 and 3 and elsewhere). While these mechanisms were not discussed using the terminology provided by Chesson (1994, 2000b) to clarify mechanisms of coexistence, they include examples of all of the types of mechanisms identified by Chesson as either slowing competitive exclusion (“equalizing mechanisms”), or promoting stable long-term coexistence (“stabilizing mechanisms”). For example, both Connell and Huston discussed the importance of compensatory (or frequency-dependent) mortality, classified by Chesson (2000b:349) as a “fluctuation-dependent stable coexistence mechanism.” Both Connell and Huston noted that increased similarity of species would slow the process of competitive exclusion, a mechanism classified by Chesson (2000b:347) as equalizing, but not stabilizing with regard to stable long-term coexistence. Perhaps most importantly, both Connell and Huston included examples of competitive nonlinearities in response to environmental fluctuations caused by disturbances, classified by Chesson (2000b:349) as “fluctuation-dependent stable coexistence mechanisms.” Trade-offs of this type, in the form of inverse r–K relationships in the species parameters, were explicitly included in the simple Lotka-Volterra competition simulations used to illustrate the DEM, and discussed in the original description of the DEM and IDH (Sheil and Burslem 2013). Model simulations demonstrate that these types of trade-offs can lead to stable coexistence under specific ranges of disturbance conditions (Roxburgh et al. 2004, Miller et al. 2011).

Disturbances that affect the most abundant species more than rare species have long been known to promote coexistence (e.g., Paine 1966). The simple density-independent disturbances used in the original DEM simulation model were selected specifically because they were the type of disturbance least likely to allow coexistence (Huston 1979:85). The other mechanisms identified by Fox as being missing from the IDH and DEM, such as temporal variations in growth rates and frequency-dependent competitive interactions (Fox 2012: Boxes 2 and 3) were well known in the literature at that time and were explicitly discussed in the context of the DEM as compensatory mechanisms that could further promote coexistence when low growth rates or disturbances slowed the process of competitive exclusion (Huston 1979:88).

The fundamental distortion introduced by Fox and also emphasized by Adler et al. is that the objective of the IDH, IPH, and DEM was to explain coexistence,
specifically the “stable long-term coexistence” that has been the focus of important theoretical work by a number of ecologists (e.g., Abrams 1983, 1995, Chesson 1985, 1994, 2000a, Chesson and Huntly 1993, 1997). However, neither Connell nor Huston emphasized, or even mentioned, stable long-term coexistence. Their focus was on predicting patterns of species diversity and explaining why the number of co-occurring species was high in some locations and low in others. Both Connell and Huston emphasized that most coexistence-promoting mechanisms can only operate effectively if rapid competitive exclusion is prevented, either by disturbances that reduce growth rates through mortality (Connell 1978, Huston 1979), or by low productivity and/or productivity–disturbance interactions that reduce growth rates and the rate of competitive exclusion (Huston 1979). The only discussions of stable coexistence in either Connell (1978) or Huston (1979) were in the context of niche partitioning and classical equilibrium competition theory, with the conclusion that, while possible under certain conditions, such coexistence was unlikely to be common in fluctuating and periodically disturbed natural environments, where most of the species required the same limited set of resources (e.g., light, water, nutrients).

Because of the central role of competitive exclusion in reducing diversity, the predictions of the IDH, IPH, and DEM only apply to groups of organisms that potentially compete with one another, such as within a single trophic level or functional type, and within an area appropriate for the organisms’ size and mobility. While different trophic levels are predicted to respond to productivity and disturbance in a manner consistent with the DEM (Huston 1994: Fig. 5.13; see Dodson et al. 2000), mixtures of trophic levels or functional types cannot be expected to behave according to the DEM (e.g., Tonkin and Death 2012).

It is clear that multiple mechanisms must operate to promote coexistence in nature. It is also clear that not all mechanisms operate everywhere or all of the time. While the simple computer model used to illustrate the predictions of the DEM did not explicitly include most of these mechanisms, it did include the two most important environmental variables for determining whether these coexistence-promoting mechanisms will have sufficient time to operate before species are eliminated by competitive exclusion. These variables are mortality and productivity (as it affects population growth rates), either of which can slow the process of competitive exclusion under specific conditions. In

---

**FIG. 3.** (A) Predictions of the dynamic equilibrium model incorporating the effects of disturbance-generated heterogeneity on the local species pool (Huston 1994: Fig. 5.10). The contour map shows variation in species diversity predicted by the dynamic equilibrium between rate of productivity (as it affects the rate of competitive displacement) and the frequency (or magnitude) of population reduction. Diversity is on the axis perpendicular to the page, and is represented by contour lines, with the highest diversity within the inner ellipsoid (in the lower left corner). The dashed-line transects demonstrate the predicted changes in diversity when one parameter is held constant and the other is varied (modified from Huston 1979: Fig. 7 and Huston 1994: Figs. 5.5 and 5.10). (B) Results of an experiment on the effects of disturbance and productivity on microbial diversity. The contour map shows diversity of colony forms of *Pseudomonas fluorescens* that developed in spatially structured (static) cultures from an initially isogenic state over a 16-d period in a multifactorial experiment with five frequencies of disturbance (every 1, 2, 4, 8, or 16 days) and eight levels of productivity, produced by a 1000× serial dilution of nutrient concentration (8× to 0.00758× normal solution). Data were transformed by Loess smoothing using a sampling proportion of 0.5. High-diversity regions are shown in lighter shades and low-diversity regions in darker shades, with diversity estimated as $1 - \sum p_i^2$, where $p_i$ is the frequency of each colony type counted from 100 plated colonies (modified from Fig. 1 in Kassen et al. 2004).
addition to the within-patch maintenance of species diversity when competitive exclusion is slowed by low growth rates or disturbances during plant succession, the spatial heterogeneity caused by multiple patches at different stages of succession (i.e., recovery from disturbance) provides spatial heterogeneity that can maintain a larger local species pool under certain combinations of productivity and disturbance, specifically where both disturbance frequency and productivity are low (Huston 1994: Fig. 5.10); see Fig. 3A.

The DEM has been implemented computationally in a variety of different ways, including individual-based plant competition models that simulate the growth and survival of hundreds of individuals of multiple species, with each individual of each species potentially experiencing different environmental conditions depending on its size, the identity of its neighbors, and the local environmental conditions (e.g., Botkin et al. 1972, Huston and DeAngelis 1994, Huston 1985, Huston and Smith 1987, Huston and Smith 1989), as well as cellular automata (Kondoh 2001). The coexistence of species, whether long-term or short-term, is ultimately based on the co-occurrence of individuals of a wide range of sizes potentially growing under a wide range of environmental conditions. These models include many of the coexistence mechanisms based on differences between species and variation in environmental conditions highlighted by Fox (2012), and demonstrate that different combinations of species characteristics can generate a wide variety of temporal population dynamics and patterns of competitive exclusion and coexistence (e.g., Huston and satisfaction 1987: Fig. 2, Huston 1994: Figs. 7.8–7.12). The various implementations of the DEM consistently produce the patterns of species diversity in relation to the interaction of productivity and disturbance predicted by the original DEM (Huston 1979). More significantly, most of the experimental efforts to test the effects of productivity and disturbance on diversity have found the predicted interaction between productivity and disturbance in relation to the conditions where species diversity is highest (Table 1, Fig. 3).

Co-occurrence vs. Coexistence

Fox (2012:87) tells us that “logically invalid hypotheses cannot hold in nature,” which is his justification for abandoning the IDH, and by inference, the DEM. However, the reality is that the DEM, and under specific conditions the IDH and IPH, does hold in nature. The DEM identified combinations of environmental conditions that are consistently associated with high species diversity without explicitly meeting the criteria for stable long-term coexistence of competitors (e.g., Abrams 1995, Chesson and Huntly 1997, Chesson 2000h).

This raises the question of whether the specific criteria that are logically required for stable long-term coexistence are actually relevant for understanding the current distribution of species diversity at some, or any, spatial or temporal scales on our planet. How can a “logically invalid hypothesis” successfully predict the observed spatial patterns and temporal dynamics of species diversity?

The simplest way that valid logic can fail to make accurate predictions is to begin with incorrect assumptions. There may be nothing wrong with the logic, but rather the problem is with the assumptions to which the logic is applied. The incorrect assumption in this case may be that stable, long-term coexistence based on competitive equilibrium is the primary determinant of the number of species that actually co-occur in a specific location. Theoretical models of equilibrium competition between species, whether or not they explicitly consider the resources for which competition occurs, make assumptions that may not be valid under some, or perhaps even most, environmental conditions. Many of these models assume that all of the competing individuals of all species have complete access to the resources for which they are competing (e.g., Tilman 1976, 1982, Tilman and Pacala 1993). This allows some species, the “best competitors,” to reduce the level of one or more resources to concentrations below that which other species require to survive. When this occurs, the poorer competitors become locally extinct and the best competitors become more abundant. This reduces species diversity to a level lower than would occur if the species were able to coexist.

The requirement that all individual organisms have equal access to the resource pool is likely to be met for microorganisms such as algae or bacteria in mixed aqueous media, such as parts of lakes under certain conditions, or laboratory chemostats. However, this assumption is much less likely to be met for plants that are permanently rooted in the soil, or for other organisms that do not predictably encounter one another in their search for resources. Analysis of models that varied the relative access of individuals to the resource pool by manipulating simulated diffusion rates found that reduced access to the resource pool potentially allowed a large number of species to co-occur regardless of their competitive abilities (Huston and DeAngelis 1994).

It seems reasonable to envision the world as containing a continuum of environments ranging from almost completely mixed with high diffusion rates (where the criteria for coexistence under equilibrium resource competition conditions apply) to unmixed with low diffusion rates (where the criteria for coexistence under equilibrium resource competition conditions are irrelevant). The ability of the DEM to predict observed patterns of species diversity in relation to productivity and disturbances suggests that some of the theoretical criteria for long-term stable coexistence may be irrelevant for understanding much of the spatial and temporal variation in species diversity on Earth.

Regardless of how, when, or if long-term stable coexistence occurs in nature at spatial and temporal
scales relevant to observed patterns of species diversity, the focus of the IDH, IPH, and DEM was originally, and is still, on predicting actual patterns of species diversity, specifically when and where diversity should be relatively high or low. Both Connell (1978) and Huston (1979) emphasized that preventing rapid competitive exclusion allowed the operation of various mechanisms that could further delay competitive exclusion. Their goal was to explain, or at least predict, patterns of species diversity that could be measured, not to demonstrate the occurrence of stable, long-term coexistence.

This raises an essential question that must be addressed before any science, either empirical or theoretical, can progress. How can theoretical predictions based on precise mathematical models be tested, or more importantly, can such theoretical predictions be tested? The empirical predictions of the IDH, IPH, and DEM can be simply tested by counting the relative number of species co-occurring under specific conditions at specific times and locations. However, it is less clear how the importance, or even the occurrence of theoretically predicted stable, long-term coexistence can be demonstrated empirically, let alone verification of the hypothesized mechanisms. Can we assume that all co-occurring species are coexisting in a manner that is stable over the long term, or are some coexisting stably and others drifting toward local extinction (or local dominance)? Over how long a time period must we make measurements of the population dynamics and competitive interactions of multiple species (hundreds of species and thousands of individuals in many plant communities) in order to determine whether locally high diversity results from coexistence that is stable and long-term or simply from slow and varying rates of local competitive exclusion?

It seems obvious that it is virtually impossible to determine whether the multiple species currently co-occurring at any specific place and time represent stable long-term coexistence or the results of other processes, such as slow rates of competitive exclusion. Consequently, the precise predictions of theoretical models of stabilizing combinations of traits among species are almost impossible to test, particularly under natural conditions, and over the spatial and temporal scales (largely unknown and unmeasured) at which stable coexistence could theoretically occur.

No ecologist has done more to clarify and summarize the relative properties and effects of various mechanisms that can potentially contribute to long-term stable coexistence than Peter Chesson (1985, 1990, 1994, 1997, 2000a, b, Chesson and Huntly 1993, 1997, Angert et al. 2009). Chesson (2000b:358) provides the following comments, which are a useful summary of some of these issues. “Other approaches to unstable coexistence... have not recognized that stabilizing components are difficult to avoid (Cohon 1997), and may have overestimated the effectiveness of purported equalizing mechanisms (Cohon 1997). Nevertheless, there is undeniable merit in the question of unstable coexistence because it must be that in many systems at least some species are only weakly persistent because their fitness disadvantages are comparable in magnitude to the stabilizing component of their long-term low-density growth rate. At this point, the study of diversity maintenance needs to take account of macroevolutionary issues such as speciation and extinction processes (Cohon and Case 1986, Hubbell 1997), biogeographic processes of migration of species between communities on large spatial scales, and climate change on large temporal scales (Chesson and Huntly 1993). Hubbell (1997) argued that on such large scales, speciation, extinction, and migration processes are dominant, rendering the admitted oversimplifications of his neutral model unimportant. Independent data on the rates of these critical processes are needed to test this perspective.”

CONCLUSIONS

The causes of high diversity clearly go beyond the simple effects of low growth rates or disturbances slowing the process of competitive exclusion, and must include multiple ecological and evolutionary processes that operate more effectively where growth rates are low and competitive exclusion occurs slowly than where growth rates are high and competitive equilibrium is reached rapidly. Fortunately, as Chesson (2000b:358) noted, “stabilizing components are difficult to avoid.” These stabilizing components include the various mechanisms of nonadditive and nonlinear competitive and growth responses that were discussed by Connell (1978) and Huston (1979, 1994) as enhancing the effects of low growth rates and disturbance on slowing competitive exclusion and allowing prolonged (although not necessarily stable, long-term) coexistence. These mechanisms must also include many of the ecological and evolutionary processes that are not included in either the DEM or simple equilibrium models of competition, such as spatial heterogeneity created by geological processes and topography, the random components of organismal dispersal and gene flow, dormant life stages such as seeds that persist through unfavorable conditions, climatic cycles, and weather patterns that dramatically alter the conditions under which organisms must survive. These and many other conditions can facilitate and prolong coexistence, regardless of whether the criteria for stable long-term coexistence are ever met.

In spite of the near impossibility of demonstrating the occurrence of long-term stable coexistence at relevant spatial and temporal scales in the real world, there are nonetheless clear patterns of species diversity measurable at scales of meters to hectares that can be predicted
based on variation in productivity and mortality. As Chesson (2000b) observed, much more work needs to be done to understand the many mechanisms that contribute to coexistence over any temporal scale, as well as those that contribute to the formation of new species. Further advances in our understanding of the regulation of diversity will come from figuring out how the many potential mechanisms that could contribute to either coexistence or speciation vary in their effects under different combinations of productivity and mortality.

Acknowledgments

I thank everyone who has found any of my ideas interesting or useful, particularly those who have used them in their own research or tested them experimentally. While I have not met most of you, I want to specifically thank Phil Regal, Norm Ellstrand, Jason Fridley, and Etienne Laliberté for their thoughtful interest. I am grateful to two anonymous reviewers who provided helpful comments on the manuscript, to Scott Wilson, who provided me with his original data from Cedar Creek, to Simon Pierce, who provided me with an earlier version of one of his published figures, and to J. W. Fox for pointing out the important paper by Kassen et al. (2004). This work was supported in part by NSF OPUS Grant 0918927 and a Developmental Leave Grant from Texas State University.

Literature Cited


I thank everyone who has found any of my ideas interesting or useful, particularly those who have used them in their own research or tested them experimentally. While I have not met most of you, I want to specifically thank Phil Regal, Norm Ellstrand, Jason Fridley, and Etienne Laliberté for their thoughtful interest. I am grateful to two anonymous reviewers who provided helpful comments on the manuscript, to Scott Wilson, who provided me with his original data from Cedar Creek, to Simon Pierce, who provided me with an earlier version of one of his published figures, and to J. W. Fox for pointing out the important paper by Kassen et al. (2004). This work was supported in part by NSF OPUS Grant 0918927 and a Developmental Leave Grant from Texas State University.
mass: the importance of environmental variables. Oikos 70:271–279.


