

DISENTANGLING THE IMPACTS OF DIVERSITY ON ECOSYSTEM FUNCTIONING IN COMBINATORIAL EXPERIMENTS

SHAHID NAEEM¹

Department of Zoology, University of Washington, 24 Kincaid Hall, Seattle, WA 98195 USA

Abstract. Biodiversity includes both taxonomic and functional aspects, each of which can play significant roles in ecosystem functioning. The number of functional groups, specifically intratrophic group (e.g., plant or herbivore) functional richness (F), serves as a simple index of ecological diversity, while species richness (S) serves as a simple index of taxonomic diversity. F and S are, however, roughly correlated measures of biodiversity, and disentangling the relative influence of one over the other on ecosystem functioning (H) requires a multivariate index. Appropriate multivariate biodiversity indices can be derived by applying principal component analyses to the set of possible combinations of S and F in an experimental design. The first principal component (PCAI) represents covariation between F and S , while the second principal component (PCAI) provides information on functioning that is associated with the independent effects of F and S . Thus, one can replace the conventional model $H = f(F, S)$ with $H = f(\text{PCAI}, \text{PCAI})$. This approach obviates a number of statistical problems encountered when following the traditional approach. Furthermore, if the question being addressed concerns the relationship between biodiversity and ecosystem functioning and not the relative contributions of F and S , PCAI may be used to develop more tractable, yet effective experimental designs than the conventional, exhaustive $F \times S$ experimental studies currently in favor. I explore the theoretical foundation for this multivariate approach and provide an example using the results from experimental prairie grassland plant assemblages at Cedar Creek Natural History Area, Minnesota, USA. This study highlights the need to adapt traditional, taxonomic approaches to biodiversity research to include functional diversity.

Key words: biodiversity; ecosystem functioning; functional groups; principal components analysis; species richness.

INTRODUCTION

Research on the relationship between diversity and ecosystem functioning has generated a new kind of experiment: one in which biodiversity is manipulated in a combinatorial fashion and ecosystem functioning is measured as a response variable to these manipulations (Naeem et al. 1995, 1996, Tilman et al. 1996, Hector et al. 1999). While the experiments themselves are relatively straightforward in design, interpreting their findings has been difficult and a source of much debate (Tilman et al. 1997b, Wardle et al. 1997, 2000, Allison 1999, Kaiser 2000, Naeem 2000). One of the most active areas in this research concerns how to determine which of several mechanistic explanations of diversity's impacts on ecosystem functioning is best supported by experimental results (Huston 1997, Sankaran and McNaughton 1999, Engelhardt and Ritchie 2001, Loreau and Hector 2001, Loreau et al. 2001, Tilman et al. 2001, Wardle 2001). Another central difficulty, which is the topic of this study, concerns disentangling the effects due to intratrophic functional diversity (e.g., functional diversity within trophic

groups such as plant functional or decomposer functional diversity) from those due to taxonomic diversity (Körner 1993, Chapin et al. 1996, Mooney et al. 1996, Gitay and Noble 1997, Hooper and Vitousek 1997, Tilman et al. 1997a, Hooper 1998, Symstad 2000, Reich and Bolstad 2001, Reich et al. 2001b). That is, ecosystem functioning (H), such as production or nutrient flux, is a function of species richness (S) and functional richness (F), or $H = f(F, S)$. Determining the relative contributions of F and S to H is important because it provides insights into the mechanisms by which biodiversity may contribute to ecosystem functioning (Hooper 1998, Hooper and Vitousek 1998). The relative contributions of F and S to H is also important in management and conservation because it provides information on ecological redundancy (Walker 1992, 1995, Lawton and Brown 1993, Gitay et al. 1996, Naeem 1998) where "ecological redundancy" refers to taxonomically different species that exhibit similar or related ecological functions (e.g., late-season grasses in plant communities or nitrogen-fixing microbes in microbial communities). Here, I focus on analytical and interpretive issues surrounding these synthetic experiments. Broader issues concerning functional groups are reviewed in Chapin et al. (1996), Mooney et al. (1996), Diaz and Cabido (2001), and Hooper et al. (*in press*).

Manuscript received 7 June 2001; revised 5 February 2002; accepted 12 February 2002; final version received 21 March 2002.

¹ E-mail: naeems@u.washington.edu

A fundamental problem in disentangling effects due to F vs. effects due to S is that F and S covary, requiring the separation of effects due to covariation from the independent effects of F and S . For most ecosystems, it is generally unknown to what degree S and F covary or are correlated. At an elementary level, they are obviously correlated in the sense that monocultures ($S = 1$) generally correspond with monofunctional groups ($F = 1$), while polycultures ($S > 1$) are very likely to have more than one functional group ($F > 1$). Thus, correlations between S and F can theoretically range from slightly above 0.0 to 1.0, depending on how tightly coupled taxonomic and functional diversity are in a given community. Currently, taxonomic diversity is more often better known or more readily derived than functional diversity, but if S and F are not strongly correlated, S will not serve as a surrogate for F , and it becomes impossible to know what the consequences to ecological function will be if the only information to hand is S . Ecologists have tackled this problem with experiments that follow factorial designs in which S and F are treated as independent factors and manipulated by varying the numbers (levels) of functional groups and species in each treatment (specific combination of S and F). Treatments are replicated and species assigned at random to each replicate in an attempt to explore the realm of all possible combinations. In theory, each treatment represents a point in a bivariate space defined by $F \times S$, and measurements of H represent points on an ecosystem functioning response surface defined by $F \times S$. One can also conceptually envision this as a vector field in which ΔH is the resultant vector, and ΔF and ΔS are the independent (orthogonal) vectors.

Unfortunately, this $H = f(F, S)$ response surface is not rectangular (Fig. 1) due to “forbidden” combinations that leave corners of the response surface empty, generating a correlation ($r_{F,S}$). Here, “forbidden” refers to combinations that either cannot be made due to the experimenter’s functional group designations (i.e., a single species belongs to one functional group, thus no combination can have $S > F$) or the limitations of experimental design do not permit certain combinations (i.e., the upper limit to S and F are set by practical limitations) (see *Designs of biodiversity experiments: Conventional designs of biodiversity . . .*, below, for further explanation). This non-rectangular (i.e., non-orthogonal) pattern created by the forbidden combinations is the root of the difficulty in separating S and F effects from one another. That is, S and F are not orthogonal, independent factors, which means that the $H = f(F, S)$ response surface is inappropriate for exploring the relative contributions of F and S to H .

While the problem of “forbidden” treatments is unavoidable, a solution to the non-orthogonal axes is readily achieved. The solution involves use of principal components (PC), a method of factor analysis in which orthogonal, multivariate axes are de-

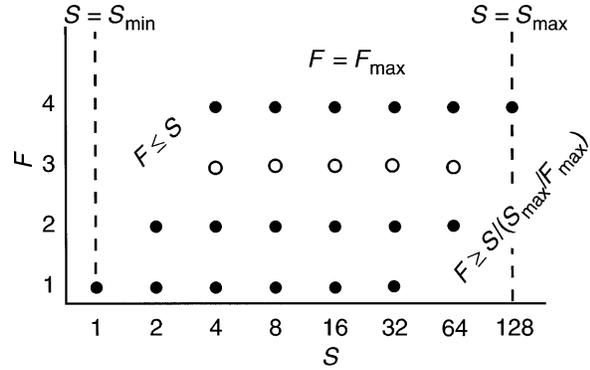


FIG. 1. The relationship between S and F in biodiversity–ecosystem functioning experimental designs. Each point represents a permissible treatment in the $S \times F$ plane. The boundaries are indicated by the formulae. Open circles represent types not usually employed because S/F is not an integer. Note the rhomboidal shape of the space indicated by the points. This rhomboid indicates that S and F are correlated. Note also that one could eliminate this correlation by designing an experiment by selecting S_{\min} and S_{\max} such that they enclose a square of points. Because the number of functional groups is generally small and near or above the often observed asymptotic relationship between S and ecosystem functioning, it is an option that is avoided.

rived from a set of covarying or correlated variables. Applying principal components analysis (PCA) to the set of S and F values for each treatment that can be employed in a full biodiversity–ecosystem functioning design provides orthogonal (completely uncorrelated) multivariate axes that provide an appropriate space for exploring the $H = f(F, S)$ response surface for the purposes of disentangling F and S effects. PCA provides an effective means for resolving the problems of disentangling F from S effects. I describe this PCA method, provide an example, and discuss the implications of this approach. The example study is taken from an experimental manipulation of plant functional group and species richness in experimental prairie grassland plots at Cedar Creek Natural History Area, Minnesota, USA.

DESIGN OF BIODIVERSITY EXPERIMENTS

Conventional designs of biodiversity–ecosystem functioning experiments

Establishing an experiment designed to test basic hypotheses about the relationship between intratrophic biodiversity and ecosystem functioning is straightforward, but difficult because of the large number of replicates needed (Naeem et al. 1995). Because the response surface is defined by $F \times S$, one would expect the number of treatments (number of levels of $F \times$ number of levels of S) to rise rapidly (e.g., $N = F \times S$, where N is the number of treatments) with increasing biodiversity. In reality, however, there are two bounds on the possible combinations of S and F . These are (1)

TABLE 1. Correlation coefficients ($r_{F,S}$) for experimental designs ranging from $S_{\max} = 1$ to 25, and $F_{\max} = 1$ to 10.

F	S								
	1	2	4	8	16	32	64	128	256
1	0	0	0	0	0	0	0	0	0
2	0	1	0.721	0.57	0.497	0.461	0.444	0.435	0.431
3	0	0	0.944	0.733	0.608	0.535	0.502	0.483	0.475
4	0	0	1	0.818	0.656	0.567	0.524	0.502	0.491
5	0	0	0	0.909	0.729	0.606	0.546	0.517	0.502
6	0	0	0	0.958	0.766	0.632	0.561	0.526	0.508
7	0	0	0	0.992	0.818	0.662	0.578	0.535	0.514
8	0	0	0	1	0.85	0.682	0.588	0.541	0.518
9	0	0	0	0	0.89	0.713	0.605	0.55	0.522
10	0	0	0	0	0.921	0.736	0.616	0.555	0.526

Note: 0 means either there is no variation in either S or F or this is a forbidden combination.

$F \leq S$, and (2) $F \geq S/(F_{\max}/S_{\max})$. Note that condition 1 is always true, while condition 2 is a convention. These bounds create a “rhomboidal” response surface as shown in Fig. 1.

It is immediately apparent from Fig. 1 that F and S are correlated. That is, using the formula for the standard Pearson product-moment correlation,

$$r_{F,S} = \frac{\sum_{i=1}^N F_i S_i - \frac{\sum_{i=1}^N F_i \sum_{i=1}^N S_i}{N}}{\sqrt{\left[\sum_{i=1}^N F_i^2 - \frac{\left(\sum_{i=1}^N F\right)^2}{N} \right] \left[\sum_{i=1}^N S_i^2 - \frac{\left(\sum_{i=1}^N S\right)^2}{N} \right]}} > 0 \tag{1}$$

where N is the number of treatments, and i is the i th F by S treatment. Using the above formula, I calculated the correlation for all experimental designs ranging from $F = 1$ to 10 and $S = 1$ to 256. Table 1 summarizes these results, which are graphed in Fig. 2. Note that statistical interpretation requires that the data meet the usual assumptions of parametric statistics, which in this case would best be met if the number of replicates for each S and F combination were the same to assure even sampling of the distribution. Furthermore, it is probably most useful when all monocultures are represented.

Because the realm of all combinations is enormous, two arbitrary conventions are usually employed in biodiversity–ecosystem functioning experiments to ensure unbiased, uniform exploration of this realm. First, researchers use levels of S as increments of 2^x , where x varies from 0 (monocultures) to $x = \ln(S_{\max})/\ln(2)$ or $\log_2(S_{\max})$. This convention reduces the number of treatments, making a more tractable experiment, and it is consistent with the observation that many patterns in nature concerning S follow logarithmic scales (Preston 1962a, b, May 1975, Rosenzweig 1995). I will henceforth refer to this practice as the “ $\log_2(S)$ convention.” Second, researchers prefer that $(S/F) = \text{an integer}$. This

convention is followed to ensure evenness of functional groups in treatments. It also permits more tractable design, substantially reducing the number of replicates. For example, following the $\log_2(S)$ convention, an $F = 4$, $S = 16$ experiment in which S/F is not constrained to integer values would require 34 treatments, whereas imposing an $S/F = \text{integer}$ constraint reduces the experimental design to just 11 treatments. Such benefits are greater for larger experiments. For example, following the $\log_2(S)$ convention, $F = 4$ and $S = 32$ makes 78 treatments, but imposing the $S/F = \text{integer}$ constraint results in a design with only 15 treatments. I will henceforth refer to this practice as the “ $S/F = \text{integer}$ ” convention.

There are several different conventional, parametric, and analytical approaches one can apply to results from the conventional biodiversity–ecosystem functioning experiments described above. These are (a) complete factorial analysis of variance (ANOVA) ($H = \text{constant} + S + F + S \times F$, where F and S are categorical variables), (b) multiple regression ($H = \text{constant} + S + F$, where F and S are treated as continuous variables), (c) analysis of covariance, model I (ANCOVA I), where F is a categorical variable and S is a covariate, and (d) ANCOVA II, where S is a categorical variable and F is a covariate.

There are several disadvantages to these conventional approaches (a–d). First, because S and F are ordered variables, information is lost if they are used as categorical variables in ANOVAs or ANCOVAs. Second, the design is unbalanced (empty cells) because of the “forbidden” treatments, which means that hypothesis testing is sensitive to the method chosen for calculating sums of squares when using general linear models. Third, because S and F are correlated, as shown above, they are not truly independent factors, making it unwise to treat them as independent categorical variables in ANOVA. Fourth, collinearity, such as that between F and S , is inappropriate for multiple regressions (Philippi 1993). Here, “collinearity” refers to one independent variable being a linear combination of other independent variables in a multiple regression,

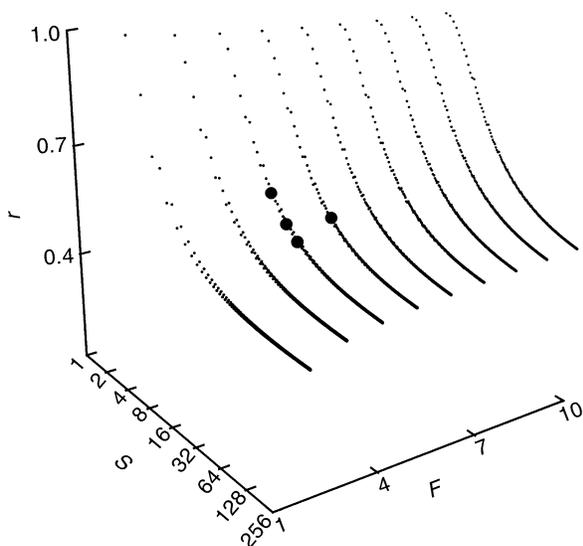


FIG. 2. Correlations between S and F ($r_{F,S}$) for experimental designs in which F ranges from 1 to 10 and S ranges from 1 to 256. These correlations are calculated for full experimental designs that do not follow the $\log_2(S)$ or $(S/F) = \text{integer}$ conventions, which represent the best experimental design possible. Correlations would be slightly stronger (an undesired result) if $\log_2(S)$ were used, and slightly weaker (a desired result) in some instances if the $(S/F) = \text{integer}$ convention were used, but the pattern would be identical. That is, the pattern showing that $r_{F,S}$ begins to level off at 0.40 would remain the same. This asymptotic approach to 0.40 indicates that it is virtually impossible to get away from the undesired correlation between S and F (a non-orthogonal design) in such $S \times F$ biodiversity-functioning experiments. Large, solid circles represent designs of three current experiments (BioCON, BIODDEPTH [e.g., Hector et al. 1999], and Cedar Creek experiments [e.g., Tilman et al. 1996, 1997]) for reference.

which inflates variation and adversely affects proper estimates of regression coefficients (Philippi 1993). Note, however, that the test remains appropriate (i.e., rejecting the null hypotheses of slope or intercept = 0.0), even if estimating coefficients is problematic.

Using PCA to identify independent F and S effects

If we consider each treatment as a measure of potential F and S influence over H , then applying PCA to this set of numbers provides two orthogonal axes that represent linear transformations of F and S (see Pielou 1984 and Manly 1994 for discussions of PCA). One may use the covariance or correlation matrix when conducting PCA, the correlation matrix being preferred when variables are unrelated (recall that correlations represent standardized covariances) (Manly 1994). PCA applied in this way identifies two axes. The first axis (PCAI) maximizes the variance it can explain and represents $r_{F,S}$. The second axis (PCAII) is orthogonal to the first and represents the remaining variance. The data can now be described by this transformation, and its values are known as PC scores. The linear transformation for PCAI is

$$S_{1,i} = \left(\frac{F_{1,i} - \bar{F}}{\sigma_F} \right) \left(\frac{c_{1,F}}{c_{1,S}} \right) \sigma_S + \bar{S} \quad (2)$$

where $c_{1,F}$ and $c_{1,S}$ are the eigenvectors for the first PCA, \bar{F} , σ_F = the mean and standard deviation of F , \bar{S} , σ_S are the mean and standard deviation for S , and $S_{1,i}$ is the transformed value for the i th S value. Note that when using the correlation matrix rather than the covariance matrix, $c_{1,F} = c_{1,S}$, so the ratio of these is 1. A similar formula can be written for $S_{2,i}$, $F_{1,i}$, and $F_{2,i}$.

After employing PC, further analyses can be done by linear regression methods. One regresses H against PCAI and PCAII using a multiple regression model. There are several advantages of this method over methods a–d described above. First, PCAI and PCAII are uncorrelated (mathematically, PCAs are orthogonal to one another) obviating the problem of collinearity between F and S . Second, PCAI provides a meaningful measure of biodiversity in the sense that it simultaneously captures taxonomic and functional diversity. Third, PCAII has the potential to provide information on the independent effects of F and S . Of course, method b, above, and this method are identical in terms of identifying a linear association between H and F and S . PCA axes represent linear transformations of the original $F \times S$ axes; therefore overall significance (P value) and the coefficient of determination (R^2 , or the ratio of sums of squares explained by regression to the total sums of squares) are identical for $H = f(S, F)$ and $H = f(\text{PCAI and PCAII})$. But because the variables used in the PCA multiple regression method are orthogonal, they are more readily interpretable than the conventional method.

AN EXAMPLE

To illustrate the above with an example, I used data from the BioCON experiment. I only briefly describe the experiment itself as complete details on the methods are published elsewhere (Reich et al. 2001a, b), and this example is provided solely for the purposes of illustration.

Methods

Briefly, this experiment closely follows the traditional biodiversity–ecosystem functioning design. It was designed to compare the impacts of plant biodiversity, elevated CO_2 , and increased rates of N deposition on ecosystem functioning (hence the acronym BioCON). I focused on the ambient plots, and thus will not provide detail on the N and CO_2 treatments. The experiment was done at the Cedar Creek Natural History Area, Anoka and Isanti Counties, Minnesota, USA. Overall, 372 plots (2×2 m) were established in 1997, of which 296 are used in the BioCON experiment. The data used here were collected in August 1999. Although many measures of ecosystem functioning were made, I restricted my analyses to percent cover, a frequently used, nondestructive, proxy mea-

TABLE 2. Plant species in the 16-species BioCON grassland plots that were used in this study.

Functional group	Plant species	Common name
C ₄	<i>Andropogon gerardii</i>	big bluestem
	<i>Bouteloua gracilis</i>	blue grama
	<i>Schizachyrium scoparium</i>	little bluestem
	<i>Sorghastrum nutans</i>	Indian grass
C ₃	<i>Agropyron repens</i>	quack grass
	<i>Bromus inermis</i>	smooth brome
	<i>Koeleria cristata</i>	Junegrass
	<i>Poa pratensis</i>	Kentucky bluegrass
Forbs	<i>Achillea millefolium</i>	yarrow
	<i>Anemone cylindrica</i>	candle anemone
	<i>Asclepias tuberosa</i>	butterfly milkweed
	<i>Solidago rigida</i>	rigid goldenrod
Legumes	<i>Amorpha canescens</i>	lead plant
	<i>Lespedeza capitata</i>	roundhead bush clover
	<i>Lupinus perennis</i>	lupine
	<i>Petalostemum villosum</i>	silky prairie clover

sure of relative aboveground plant biomass production. Percent cover was estimated evaluating percent cover of each species in 100 × 50 cm frames placed in identical positions in each plot, one frame per plot.

The ground was cleared and treated with methyl bromide to remove the seedbank and weeds. Plots were sprayed with a filtered slurry to reintroduce microbes, and seeded to establish the diversity treatments. The pool of plant species and their functional group designations are listed in Table 2. *S* varied from 0, 1, 4, 9, and 16 species, while *F* varied from 0, 1, 2, 3, to 4 groups. I restricted analyses to those plots in which *S* > 0 and *F* > 0.

In the first analysis, I considered the original design and the realized design, and compared diversity in terms of relative abundance using the Shannon index (Shannon and Weaver 1949, Magurran 1988) applied to percent cover data. In the second analysis, I examined plots under ambient conditions to determine how percent cover responded to variation in *F* and *S* using both conventional methods and the PC method described above.

Results

For the BioCON example here, the original design, following the $\ln_2(S)$ convention, with $S_{\min} = 1$, $S_{\max} = 16$, $F_{\min} = 1$, and $F_{\max} = 4$, produces 11 possible community types (Fig. 3).

By comparison, only seven combinations were used in the actual experimental design, but this design still yields an $r_{F,S}$ of 0.661 (Fig. 4), which is only slightly higher than the value for the full design. Ordinarily, 1, 2, 4, 8, and 16 species would be used, but the researchers selected a subset of these possible levels, using 1, 4, 9, and 16 species instead. This closely follows the $\log_2(S)$ convention. The actual design also deviated slightly from the $S/F = \text{integer}$ convention in that $S = 9$ does not yield an integer when divided by $F = 4$, but densities of seeds were still apportioned equally to ensure even distribution of functional groups. Because the levels of *S* (1, 4, 9, 16) are relatively evenly spread across the range, I did not employ $\log_2(S)$ transformations in further analyses.

On average, the actual BioCON design is very close to a full, traditional design, even though there are fewer treatments (Table 3). The smaller number of treatments did result in more species- and functionally rich plots (fewer levels for the same range) and slightly less variation in *S* and *F* (more homogeneity), but these differences are small (Table 3), indicating that the actual design has relatively the same power as a full design.

Local extinction and dominance invariably reduces

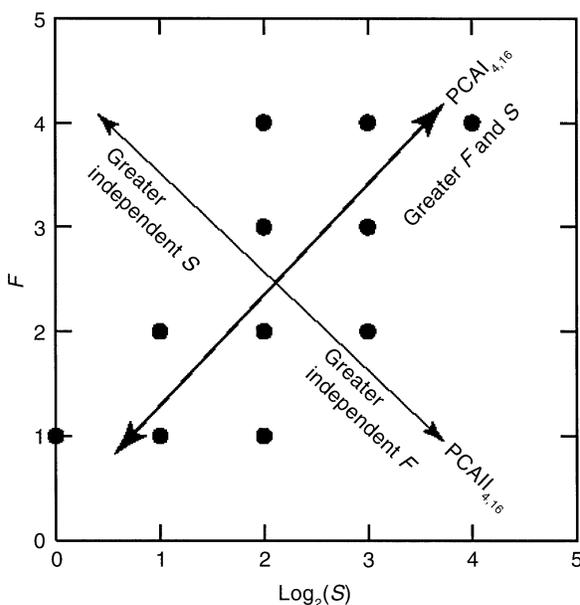


FIG. 3. Use of PCA in disentangling *F* and *S* effects for the BioCON experiment. $PCAI_{4,16}$ = principal component axis I for $F = 4$ and $S = 16$. $PCAI_{4,16}$ is similarly defined. Points indicate possible treatments.

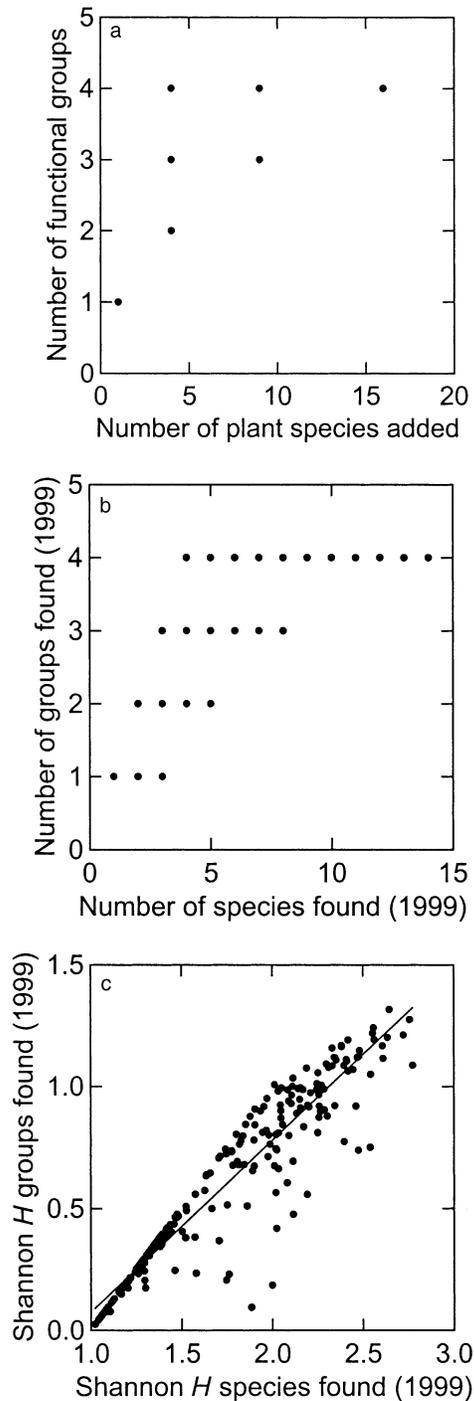


FIG. 4. Results from the example (BioCON) study. Panel (a) shows actual treatments used; (b) shows actual combinations of F (y-axis) and S (x-axis) found in plots based on percent cover quadrats, and (c) shows relationship between F and S where the Shannon formula was applied to percent cover quadrats to estimate diversity based on relative abundance actually observed in the plot.

the diversity of a plot from its initial diversity. While both F and S may not be affected, determining the extent to which they remain unaffected by succession is difficult to determine as it would require a thorough search of every plot. Percent cover quadrats, however, revealed that the pattern created by the $\log_2(S)$ convention alters substantially if extreme rarity is treated as equivalent to extinction.

The actual $F \times S$ space being explored by the experimental system contains many replicates that may have deviated from the $\ln_2(S)$ and $S/F = \text{integer}$ conventions due to local extinction or rarity. Because percent cover estimates are not capable of confirming the presence or absence of extremely rare plants, it is not certain how far from the original design the replicates deviated, but such deviations do not change the nature of the experiment, only the extent of the response surface being explored. The actual experiment may have covered the set of possible combinations of F and S more thoroughly by deviating from the two conventions. Unlike the possible reduced range of S (S_{\max} observed = 14), the range of F values explored was unaffected because at least one individual of each functional group was found in replicates where they were planted.

The Shannon index, which weighs presence by abundance, shows a tight distribution around a positive, nearly linear association between F and S (Fig. 4). Clearly, actual taxonomic diversity (S) is highly correlated with functional diversity (F) ($r = 0.949$) in this experimental system if we consider relative abundance. Applying PCA to the Shannon index of F and S as determined by percent cover, PCA I describes 97.446% and PCA II describes 2.554% of the total variance, indicating that there is even less power in this experiment to separate F and S effects if we consider relative abundance.

Different methods of conventional parametric analyses yield different results (Table 4). Total percent cover, the response variable, shows a strong positive relationship with all measures of biodiversity (F , S , and PCA I) (Fig. 5). Of particular note is the nearly identical pattern with either F or S , confirming strong collinearity. The imbalance of the design did not permit a full multifactorial ANOVA. Separate tests using general linear models revealed no significant main effects (F and S), no significant interaction ($F \times S$), and no significance if either is treated as a covariate (Table 4).

Multiple regression did yield a significant regression (Table 4), with F proving to be a significant ($P < 0.05$) contributor to the regression. One might conclude from this that functional richness is more important, but the collinearity between these two variables does not allow for unambiguous interpretation. In contrast, PCA I showed a highly significant ($P < 0.001$) contribution to the regression. Because PCA I describes the correlation between F and S , PCA II provides information on the relationship between functioning that is unaf-

TABLE 3. Summary statistics and principal component analysis (PCA) results for the full BioCON and actual BioCON design.

a) Summary statistics						
	$\bar{F}_{\log_2}(\sigma F_{\log_2})$	$\bar{S}_{\log_2}(\sigma S_{\log_2})$	$\bar{S}(\sigma)$	Levels		$r_{F,S}$
				S	F	
Full	2.46 (1.21)	2.09 (1.14)	5.55 (4.28)	1, 2, 4, 8, 16	1, 2, 3, 4	0.66
Actual	3.00 (1.16)	...	6.71 (5.02)	1, 4, 9, 16	1, 2, 3, 4	0.66

b) Principal component analysis				
	PCI	PCII	PCI	PCII
Variance explained (%)	84.61	15.39	83.04	16.96
Eigenvector				
F	0.71	0.71	0.71	0.71
S	0.71	-0.71	0.71	-0.71
Component loading				
F	0.92	0.39	0.91	0.41
S	0.92	-0.39	0.91	-0.41

Notes: The first principal component (PCI) and second principal component (PCII) = principal components 1 and 2, respectively. Note that PCA used $\log_2(S)$ for the analysis of the full design, but S for the actual design due to the approximate \log_2 spread of levels. $\bar{F}_{\log_2}(\sigma F_{\log_2})$ = average and standard deviation of F , \log_2 transformed. $\bar{S}_{\log_2}(\sigma S_{\log_2})$ = average and standard deviation of S , \log_2 transformed. $\bar{S}(\sigma)$ = average and standard deviation of untransformed S . $r_{F,S}$ = correlation between F and S . This table shows that the full design and actual design are quite similar in design properties.

ected by F and S . Only 1.3% of the variance in percent cover is explained by PCII. A positive slope (linear regression, coefficient of $F = 4.91$, 1 SE = 3.49, lower 95% CI = -2.05, upper 95% CI = 11.87, $R^2 = 0.01$, $P = 0.16$), suggests a trend in increasing production being partly associated with independent F effects, but lack of significance ($P > 0.05$), low R^2 , and the fact that the 95% confidence interval for the coefficient includes 0 does not support this hypothesis.

DISCUSSION

This study shows clearly that for any experiment examining $H = f(F, S)$, $r_{F,S}$ is unavoidable. In fact, such experiments are particularly poor at detecting possible

independent F and S effects. While larger experiments have lower correlations (Table 1, Fig. 2), even at 256 species, PCII indicates that only 33.11% of the variance in the experimental design can be used for determining independent F and S effects. While this is a gain of 17.73% over an experiment using 32 species, the increase from 15 treatments to 67 indicates that considerable effort may be necessary to make small gains in experimental power. Thus, conclusions that functional group richness matters more than species richness based on the limited set of current experiments are not strongly supported, since they lack the power to disentangle the two factors. If resources for conducting research are limited and the main goal is to

TABLE 4. Comparison of statistical methods that can be applied to results from a biodiversity ecosystem functioning.

Method	Model term or effect	df	P
ANOVA (GLM)	F (category)	2, 67	0.943
	S (category)	2, 67	0.805
	$F \times S$ (interaction)	2, 67	0.855
Multiple regression	F	2, 71	0.038
	S		0.599
	R^2		0.239
ANCOVA I	F (category)	3, 69	0.195
	S (covariate)	1, 69	0.529
ANCOVA II	S (category)	3, 69	0.787
	F (covariate)	1, 69	0.865
PCA-multiple regression	PCAI	2, 71	< 0.001
	PCAI I		0.448
	R^2		0.239

Notes: Only ambient condition plots were used ($N = 74$) in which $S > 0$ and $F > 0$. The response variable is percent vegetation cover. P = probability value where $P < 0.05$ is considered significant (boldface type).

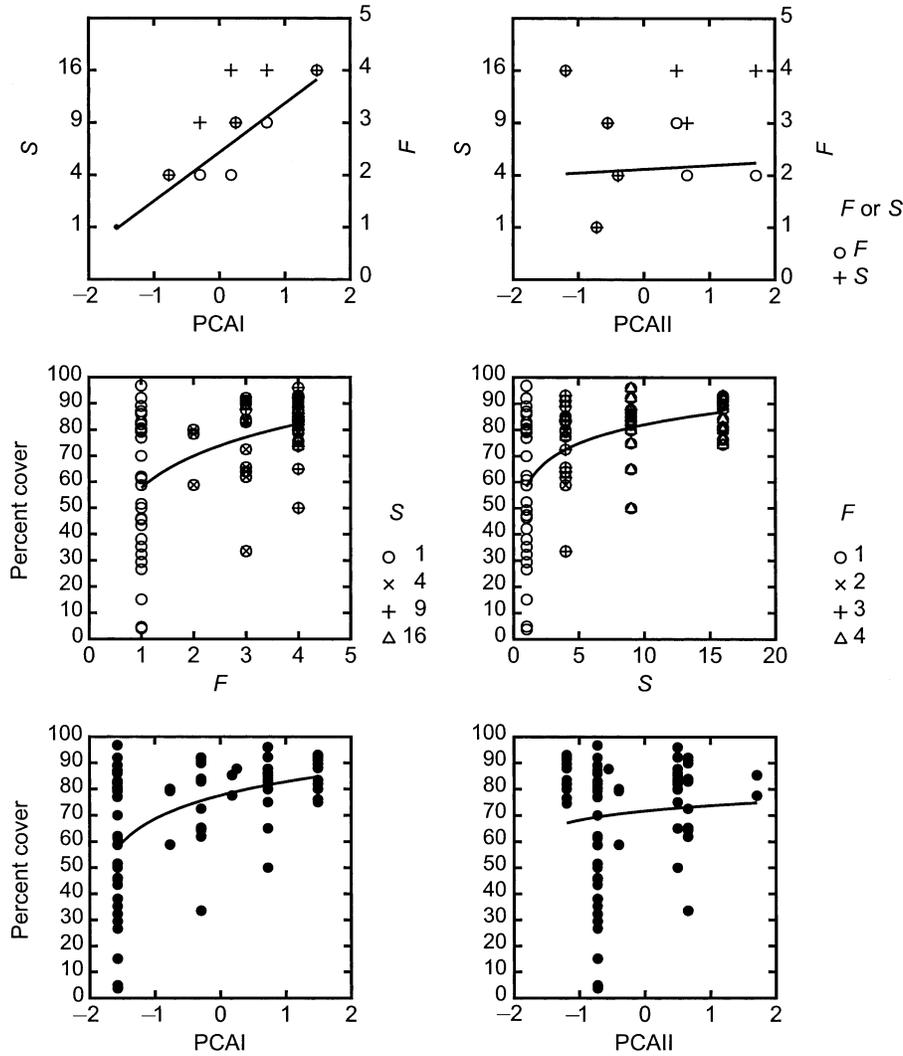


FIG. 5. Plots of ecosystem functioning (percent cover) vs. F and S , and of percent cover vs. PCAI and PCAII for the BioCON example. The top plots show relationships between F and S (right and left vertical axes, respectively) and PCAI and PCAII. Note the strong correlation identified by PCAI. The middle plots show percent cover in relation to F and S . The bottom plots show percent cover in relation to PCAI and PCAII. Lines in middle and bottom plots represent best-fit log-linear relationships.

determine the association between biodiversity and ecosystem functioning, then investing in the full experiment may be unwise.

If resources are not a limiting factor and one is interested in disentangling the relative impacts of F and S on H , however, then a complete, factorial experimental design is desirable. The response surface, however, is potentially enormous, requiring large numbers of treatments even if following the $\log_2(S)$ and $S/F =$ integer conventions for subsampling the set of F and S combinations. Given the importance of the issues being addressed by such studies, these daunting requirements should not (and have not) dissuaded experimental ecologists from conducting such studies. Applying conventional, parametric analytical methods to the results, however, is problematic if the goal is to

disentangle F and S effects. The PCA solution provided here provides means for gaining some insights into relative impacts of both covarying F and S and independent F and S effects.

There are three basic steps involved in this approach. First, one must design the experiment, which involves determining the appropriate S_{\max} , S_{\min} , F_{\max} , and F_{\min} for the ecosystem under investigation, and consider whether the $\log_2(S)$ and $S/F =$ integer conventions will be adopted. This is a critical step since results are sensitive to which species are selected and to what scheme for functional group classification is adopted. Equally important, and potentially problematic, is the fact that functional group classification schemes are often optimal for single ecosystem functions (dependent variables), thereby necessitating the use of different

schemes for widely differing ecosystem functions. Here, an optimal functional classification scheme refers to one in which the trait-group distances among species within groups is smaller than the distances among functional groups (e.g., Wardle 2001). Because trait groups used in functional classification schemes are generally associated with the dependent ecosystem function under investigation (Chapin et al. 1996), considerable caution is necessary when conducting analyses of two or more dependent ecosystem function variables, but using only one functional classification scheme.

The second step involves analyzing the experimental design using PCA as done above. The third step involves measuring H and using it as a dependent variable in a regression analysis in which PCAI and PCAII are treated as the independent variables (rather than against F and S). This approach may be extended to multiple dimensions. For example, if an experiment were to manipulate both trophic and intratrophic functional groups, a third axis may represent a number of trophic levels. Manipulations of intertrophic functional groups have similar limitations as manipulations of intratrophic groups. For example, the number of intertrophic groups must be equal to or less than the number of intratrophic groups. The resulting statistical procedure would be three regression analyses using PCAI, PCAII, and PCAIII as independent variables. PCAI, in this case, would represent the multivariate correlation among species, intratrophic, and intertrophic functional richness in the experimental design. Interpretation of PCAII and PCAIII would depend on factor loadings, but would represent the diversity effects independent of the multivariate correlation among the components of diversity.

One potentially valuable use of this method may be in designing simpler biodiversity–ecosystem functioning experiments. This analysis can be conducted before the experiment to aid in decisions concerning the magnitude and scope of the experiment. If the purpose of the experiment is to determine the relationship between biodiversity and ecosystem functioning, then one can use PCAI to determine what treatments to use. For a set of values of F , ranging from S_{\min} to S_{\max} , an i th value of F in the experimental set, we can determine S_i using the formula for PCAI (Eq. 2). Using this formula, for each level of functional (F_i) groups from 1 to F_{\max} , one could select the number of species at or near the value of S_i on PCAI. For example, if $S_{\min} = 1$, $S_{\max} = 32$, $F_{\min} = 1$, and $F_{\max} = 4$, there are 15 $S \times F$ treatment levels following the $\log_2(S)$ convention. If we replicate each one at 10 replicates per level, this yields 150 replicates. PCAI for this design, however, indicates that the set of treatments where $S \approx 2^F$ (effectively four points), covers the range of biodiversity in this experimental design. This would allow reduction of the experiment to either four treatments with 10 replicates for each treatment ($N = 40$), or, if resources permit, ~150 replicates: four treatments with 37 replicates

each. The former provides for a tractable experiment for determining a biodiversity effect as the larger design, whereas the latter allows for greater statistical power and greater ability to explore the realm of species combinations, which is often severely undersampled. For example, the number of possible unique treatments equals 6.12×10^8 for 32 species, even if following the $\log_2(S)$ convention.

There are several cautions one must still follow. First, it is important to recall that defining functional diversity is problematic (Gitay and Noble 1997). So long as functional groups are defined operationally or qualitatively, interpretation of findings is necessarily sensitive to these definitions. Any change in the definition that changes either the number of functional groups or the classification of species can change the outcome. Second, variance in these experiments is still heterogeneous, which makes using regression models suspect. One can solve this by dividing each level by its standard deviation, but such a transformation is difficult to interpret since it effectively makes $H/\sigma_{F,S} = f(F, S)$.

There are also several issues concerning functional groups that intratrophic $S \times F$ experiments and this method cannot address. While the functioning of an ecosystem is ultimately related to the functional properties of the species within its community, ecosystem functioning is likely to be associated with F or even just the identity of a species (where identity refers to the specific contribution of a species to ecosystem functioning) (Hooper and Vitousek 1997, Tilman et al. 1997a, Symstad 2000, Hooper et al., *in press*), but there are several issues that make it difficult to adopt this simple principle. First, the issue of identifying the appropriate mechanism underlying biodiversity effects (e.g., Huston 1997, Loreau and Hector 2001) is the same for functional diversity as it is for taxonomic diversity. Multiple functional groups may act through sampling, niche complementarity, facilitation, or possibly other mechanisms in similar ways that multiple species create associations between biodiversity and ecosystem functioning. Second, intertrophic functional richness may create relationships between biodiversity and functioning that are stronger than intratrophic functional diversity (Naeem and Li 1998, Mulder et al. 1999, Hulot et al. 2000, Naeem et al. 2000, Loreau 2001). Third, the predictability or variability of a system may be dependent on the number of redundant species or number of species within a functional group (McGrady-Steed et al. 1997, Naeem and Li 1997, Doak et al. 1998, Naeem 1998, Tilman et al. 1998, Yachi and Loreau 1999). Thus, even if the central tendency or mean levels of ecosystem functioning are insensitive to variation in taxonomic diversity, variability of functioning might be sensitive to S or S/F . Fourth, from a practical standpoint, in comparison to taxonomic methods, neither the knowledge of species' functional identities nor functional classification systems are devel-

oped sufficiently to adequately quantify functional biodiversity, though progress is being made (Hooper et al., *in press*). Thus, while plant experiments such as those by Hooper and Vitousek (1997), Tilman et al. (1997), Hector et al. (2000), Engelhardt and Ritchie (2001), Symstad and Tilman (2001), and Reich et al. (2001a) provide valuable insights into plant functional diversity and ecosystem functioning, they shed little light into diversity across multiple trophic groups or diversity effects on system variability, a possibility that aquatic microcosm experiments have shed some light on (Petchey et al., *in press*).

Of these many conceptual issues surrounding the relative roles of taxonomic and functional diversity in ecosystem functioning, the PCA method reported here primarily solves the non-orthogonality of the F and S variables. I suggest, however, that in many instances the relative contributions of F and S is not what motivates research addressing the ecosystem consequences of declining biodiversity. Taxonomic and functional diversity are inextricably linked in such experiments; thus there may be little to be gained in attempting to understand what the independent effects of F and S might be by this method. What is important, however, is that direct manipulations of biodiversity manipulate F and S in a balanced way, and the PCA method provides guidelines for doing this. In the interest of time and limited resources, experiments that explore the diagonal of the $F \times S$ matrix (identified by PCAI) may provide a more effective means of gaining insight into the relationship between biodiversity and ecosystem functioning. The reduction in the numbers of treatments resulting from this approach may be particularly useful when attempting to expand biodiversity studies to multitrophic levels where reduction in the number of treatments is necessary to make such experiments tractable. Such reductions in experimental design may also facilitate identifying the underlying mechanisms of biodiversity impacts on ecosystem functioning (e.g., sampling, niche complementarity, and facilitation) by permitting greater allocation of resources to replication of species combinations, one of the biggest challenges to biodiversity-functioning research (Loreau et al. 2001).

ACKNOWLEDGMENTS

I thank S. Lavorel, S. F. Tjossem, and two anonymous reviewers for critically reading the manuscript. D. Bahaudin, W. Bengston, J. Goth, and T. Mielke were instrumental in managing the BioCON experiment from which the data for the example were derived. This research was supported by grants from the National Science Foundation (NSF DEB 9734343 and NSF DEB 0080382) and the Department of Energy.

LITERATURE CITED

- Allison, G. W. 1999. The implications of experimental design for biodiversity manipulations. *American Naturalist* **153**: 26–45.
- Chapin, F. S., III, H. L. Reynolds, C. M. D'Antonio, and V. M. Eckhart. 1996. The functional role of species in terrestrial ecosystems. Pages 403–428 in B. Walker and W. Steffan, editors. *Global change and terrestrial ecosystems*. Cambridge University Press, Cambridge, UK.
- Diaz, S., and M. Cabido. 2001. Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology and Evolution* **16**:646–655.
- Doak, D. F., D. Bigger, E. Harding-Smith, M. A. Marvier, R. O'Malley, and D. Thomson. 1998. The statistical inevitability of stability-diversity relationships in community ecology. *American Naturalist* **151**:264–276.
- Engelhardt, K. A. M., and M. E. Ritchie. 2001. Effects of macrophyte species richness on wetland ecosystem functioning and services. *Nature* **411**:687–689.
- Gitay, H., and I. R. Noble. 1997. What are functional types and how should we seek them? Pages 20–46 in T. M. Smith, H. H. Shugart, and F. I. Woodward, editors. *Plant functional types*. Cambridge University Press, Cambridge, UK.
- Gitay, H., J. B. Wilson, and W. G. Lee. 1996. Species redundancy: a redundant concept? *Journal of Ecology* **84**: 121–124.
- Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Högberg, et al. 1999. Plant diversity and productivity experiments in European grasslands. *Science* **286**:1123–1127.
- Hooper, D. U. 1998. The role of complementarity and competition in ecosystem responses to variation in plant diversity. *Ecology* **79**:704–719.
- Hooper, D. U., et al. *In press*. Species diversity, functional diversity and ecosystem functioning. In M. Loreau, S. Naeem, and P. Inchausti, editors. *Biodiversity and ecosystem functioning*. Oxford University Press, Oxford, UK.
- Hooper, D. U., and P. M. Vitousek. 1997. The effects of plant composition and diversity on ecosystem processes. *Science* **277**:1302–1305.
- Hooper, D. U., and P. M. Vitousek. 1998. Effects of plant composition and diversity on nutrient cycling. *Ecological Monographs* **68**:121–149.
- Hulot, F. D., G. Lacroix, F. Lescher-Moutoué, and M. Loreau. 2000. Functional diversity governs ecosystem response to nutrient enrichment. *Nature* **405**:340–344.
- Huston, M. A. 1997. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia* **110**:449–460.
- Kaiser, J. 2000. Rift over biodiversity divides ecologists. *Science* **289**:1282–1283.
- Körner, C. 1993. Scaling from species to vegetation: the usefulness of functional groups. Pages 117–132 in E.-D. Schulze and H. A. Mooney, editors. *Biodiversity and ecosystem functioning*. Springer-Verlag, Berlin, Germany.
- Lawton, J. H., and V. K. Brown. 1993. Redundancy in ecosystems. Pages 255–270 in E.-D. Schulze and H. A. Mooney, editors. *Biodiversity and ecosystem function*. Springer-Verlag, Berlin, Germany.
- Loreau, M. 2001. Microbial diversity, producer–decomposer interactions and ecosystem processes: a theoretical model. *Proceedings of the Royal Society of London, Series B* **268**: 303–309.
- Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. *Nature* **412**: 72–76.
- Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294**:806–808.
- Magurran, A. E. 1988. *Ecological diversity and its measurement*. Princeton University Press, Princeton, New Jersey, USA.
- Manly, B. F. J. 1994. *Multivariate statistical methods*. Chapman and Hall, London, UK.

- May, R. M. 1975. Patterns of species abundance and diversity. Pages 81–120 in M. L. Cody and J. M. Diamond, editors. *Ecology and evolution of communities*. Belknap Press of Harvard University, Cambridge, Massachusetts, USA.
- McGrady-Steed, J., P. M. Harris, and P. J. Morin. 1997. Biodiversity regulates ecosystem predictability. *Nature* **390**:162–165.
- Mooney, H. A., J. H. Cushman, E. Medina, O. E. Sala, and E.-D. Schulze. 1996. *Functional roles of biodiversity: a global perspective*. John Wiley and Sons, New York, New York, USA.
- Mulder, C. P. H., J. Koricheva, K. Huss-Danell, P. Högborg, and J. Joshi. 1999. Insects affect relationships between plant species richness and ecosystem processes. *Ecology Letters* **2**:237–246.
- Naeem, S. 1998. Species redundancy and ecosystem reliability. *Conservation Biology* **12**:39–45.
- Naeem, S. 2000. Reply to Wardle et al. *ESA Bulletin* **81**:241–246.
- Naeem, S., K. Haakenson, L. J. Thompson, J. H. Lawton, and M. J. Crawley. 1996. Biodiversity and plant productivity in a model assemblage of plant species. *Oikos* **76**:259–264.
- Naeem, S., D. Hahn, and G. Schuurman. 2000. Producer-decomposer codependency modulates biodiversity effects. *Nature* **403**:762–764.
- Naeem, S., and S. Li. 1997. Biodiversity enhances ecosystem reliability. *Nature* **390**:507–509.
- Naeem, S., and S. Li. 1998. Consumer species richness and autotrophic biomass. *Ecology* **79**:2603–2615.
- Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1995. Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Philosophical Transactions of the Royal Society of London, B* **347**:249–262.
- Petchey, O. L., P. J. Morin, F. Hulot, M. Loreau, J. McGrady-Steed, G. Lacroix, and S. Naeem. *In press*. Contributions of aquatic model systems to our understanding of biodiversity and ecosystem functioning. Oxford University Press, Oxford, UK.
- Philippi, T. E. 1993. Multiple regression: herbivory. Pages 183–210 in S. M. Scheiner and J. Gurevitch, editors. *Design and analysis of ecological experiments*. Chapman and Hall, New York, New York, USA.
- Pielou, E. C. 1984. *The interpretation of ecological data: a primer of classification and ordination*. John Wiley and Sons, New York, New York, USA.
- Preston, F. W. 1962a. The canonical distribution of commonness and rarity: part I. *Ecology* **43**:185–215.
- Preston, F. W. 1962b. The canonical distribution of commonness and rarity: part II. *Ecology* **43**:410–432.
- Reich, P. B., and P. Bolstad. 2001. Productivity of evergreen and deciduous temperate forests. Pages 245–283 in J. Roy, B. Saugier, and H. A. Mooney, editors. *Terrestrial productivity*. Academic Press, San Diego, California, USA.
- Reich, P. B., J. Knops, D. Tilman, J. Craine, D. Ellsworth, M. Tjoelker, T. Lee, D. Wedin, S. Naeem, D. Bahaouddin, G. Hendrey, S. Jose, et al. 2001a. Plant diversity influences ecosystem responses to elevated CO₂ and nitrogen enrichment. *Nature* **410**:809–812.
- Reich, P. B., D. Tilman, J. Craine, D. Ellsworth, M. Tjoelker, J. Knops, D. Wedin, S. Naeem, D. Bahaouddin, J. Goth, W. Bengston, and T. Lee. 2001b. Do functional groups differ in acquisition and use of C, N, and water under varying atmospheric CO₂ and N deposition regimes? A field test using 16 grassland species from 4 functional groups. *New Phytologist* **150**:435–448.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, New York, New York, USA.
- Sankaran, M., and S. J. McNaughton. 1999. Determinants of biodiversity regulate compositional stability of communities. *Nature* **401**:691–693.
- Shannon, C., and W. Weaver. 1949. *Mathematical theory of communication*. University of Illinois Press, Urbana, Illinois, USA.
- Symstad, A. J. 2000. A test of the effects of functional group richness and composition on grassland invasibility. *Ecology* **81**:99–109.
- Symstad, A., and D. Tilman. 2001. Diversity loss, recruitment limitation, and ecosystem functioning: lessons learned from a removal experiment. *Oikos* **92**:424–435.
- Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Sieman. 1997a. The influence of functional diversity and composition on ecosystem processes. *Science* **277**:1300–1302.
- Tilman, D., C. L. Lehman, and C. E. Bristow. 1998. Diversity–stability relationships: statistical inevitability or ecological consequence? *American Naturalist* **151**:277–282.
- Tilman, D., S. Naeem, J. Knops, P. Reich, E. Siemann, D. Wedin, M. Ritchie, and J. Lawton. 1997b. Biodiversity and ecosystem properties. *Science* **278**:1866–1867.
- Tilman, D., P. B. Reich, J. Knops, D. Wedin, T. Mielke, and C. L. Lehman. 2001. Diversity and productivity in a long-term grassland experiment. *Science* **294**:843–845.
- Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* **379**:718–720.
- Walker, B. H. 1992. Biological diversity and ecological redundancy. *Conservation Biology* **6**:18–23.
- Walker, B. 1995. Conserving biological diversity through ecosystem resilience. *Conservation Biology* **9**:747–752.
- Wardle, D. A. 2001. Experimental demonstration that plant diversity reduces invasibility—evidence of a biological mechanism or a consequence of sampling effect? *Oikos* **95**:161–170.
- Wardle, D. A., M. A. Huston, J. P. Grime, F. Berendse, E. Garnier, W. K. Lauenroth, H. Setälä, and S. D. Wilson. 2000. Biodiversity and ecosystem function: an issue in ecology. *ESA Bulletin* **81**:235–239.
- Wardle, D. A., O. Zackrisson, G. Hornberg, and C. Gallet. 1997. Biodiversity and ecosystem properties. *Science* **278**:1867–1869.
- Yachi, S., and M. Loreau. 1999. Biodiversity and ecosystem functioning in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences (USA)* **96**:1463–1468.