No Consistent Effect of Plant Diversity on Productivity

Hector et al. (1) reported on BIODEPTH, a major international experiment on the response of plant productivity to variation in the number of plant species. They found “an overall log-linear reduction of average aboveground biomass with loss of species,” leading to what the accompanying Perspective (2) described as “a rule of thumb—that each halving of diversity leads to a 10 to 20% reduction in productivity.” These conclusions, if true, imply that the continuing high rate of plant extinction threatens the future productivity of Earth’s natural and managed ecosystems and could impair their ability to produce resources essential for human survival and to regulate the concentration of atmospheric CO₂.

Several problems with the Hector et al. article, however, lead us to question its major conclusions. First, the experimental approach of random species addition or removal in immature plant assemblages mimics neither natural nor human-caused processes of species extinction, accumulation, or combination and, thus, likely has little practical relevance. Second, the statistical analyses of Hector et al. made assumptions incompatible with the experimental design, thereby undermining many of the study’s conclusions. And, third, the article did not quantify crucial treatment and response variables, which reduces confidence in any conclusions about treatment effects. Other authors have thoroughly discussed the nonrandom nature of species establishment and extinction (3–5); here, we focus on issues of experimental design and statistical analysis.

The most important conclusion was that “a single general relationship” may exist between species richness and productivity (1). This conclusion rested on two statistical findings: (i) that across all eight sites there were no statistically significant differences between the patterns observed, and (ii) that within seven of the eight sites there were statistically significant differences in productivity between the diversity treatments imposed at the individual site, which led to the reported within-site regressions and analyses of variance (ANOVARs). Even if both findings were technically correct, however, they do not necessarily imply that increasing species number causes a predictable increase in plant productivity.

The data of Hector et al. do not support the conclusion that all sites showed a similar positive relationship between diversity and productivity. For three of the sites (Greece, Ireland, and Silwood), the authors’ own analysis showed no consistent change in biomass with species number, which our own regression of their data confirmed (Fig. 1). Moreover, two of the five sites at which Hector et al. identified significant positive regressions (Germany and Switzerland) included more species in their highest-diversity treatments than in the monocultures, so these treatments cannot be included in the regression without violating a crucial statistical constraint. The demonstration that plant productivity increases with species number in a mixture requires that the mixture’s productivity be greater than that of any species from the mixture grown separately, a response known as overyielding (6–11). If a mixture includes species whose growth has not been measured in a monoculture, it is impossible to determine whether the higher productivity of the mixture results from the biological processes that potentially cause overyielding or simply from the addition of a very productive species that was not evaluated in monoculture. Thus, the maximum number of species in the highest-diversity mixture must be no greater than the number of species that are grown in monoculture, and all of the highest-diversity mixtures must be identical in species composition, since each must contain only species grown in monoculture.

The species assemblage codes in the BIODEPTH dataset show that only at the Portugal, Sweden, and Sheffield sites do all replicates of the highest-diversity treatment (species number half the number of monoculture plots) have identical species composition. At the other five sites, multiple species codes in the high-diversity treatments indicate that these mixtures include species not evaluated in monoculture. Five of the eight sites, therefore, cannot legitimately be used in statistical analyses of overyielding and diversity-productivity relationships, because the analysis may be biased toward increasing the productivity of high-diversity mixtures. The inclusion of unevaluated species in mixtures may explain the high variability in the biomass responses at these five sites compared with the three sites with properly designed experiments. That high variability allowed the authors to conclude that all sites had statistically similar responses. However, the inability to distinguish “no response” from “positive response” suggests that the real treatment effects were weak.

The three sites with proper experimental design (Portugal, Sweden, and Sheffield) all showed significant positive regressions of productivity across two or three doublings of species richness [Fig. 1; (12)]. This is the pattern expected from random selection from a set of objects with different properties (13–15), because the probability of including any specific member of the set—such as a plant species that grows rapidly or fixes nitrogen—increases with the number of objects selected. Such a pattern, found consistently in randomly assembled experimental plant communities but only rarely in natural plant communities (4, 5, 13–15), has been identified as a statistical artifact of experimental design (5, 13, 14). Although one study (15) suggested that the pattern constitutes a natural mechanism by which diversity affects productivity, this requires the biologically unrealistic assumption that plant communities are randomly assembled with respect to productivity (5).

Separation of the effects of random selection from those of biological interactions that
increase productivity again requires demonstration of the phenomenon of overyielding, discussed above. The statistical evaluation of overyielding has been thoroughly discussed (6–11), yet Hector et al. used a nonstandard method that inevitably must overestimate overyielding, because it does not compare mixture productivity with the most productive monoculture [note 33 of (1)]. Indeed, notwithstanding the assertion of Hector et al. [(1), p. 1126] that the BIODEPTH experiments demonstrated that overyielding increases with increasing species richness, inspection and our reanalysis of the data clearly show that mixtures with many species are no more productive than the most productive monocultures. In our reanalysis, none of the three sites with an experimental design appropriate for analysis of overyielding showed overyielding in response to species richness (16); the near-zero slopes of the overyielding regressions for Portugal, Sweden, and Sheffield demonstrate that the multispecies mixtures were no more productive than the most productive monocultures. This is the pattern of maximum productivity expected when the increase in average productivity is a statistical artifact of random sampling (5, 13, 14).

Nonetheless, some plots clearly did show true overyielding. The strongest overyielding response in the entire set of experiments resulted from the addition of a single species of legume (Trifolium pratense), which increased productivity by an average of 360 g m\(^{-2}\)—more than the average difference between the highest- and lowest-diversity treatments at most sites, and more than four times greater than the reported result that “each halving of the number of plant species reduced productivity by approximately 80 g m\(^{-2}\) on average” [(1), p. 1124]. Thus, the primary cause of overyielding in many of the mixtures was not the total number of species, but simply the fertilization effect of nitrogen-fixing legumes.

Another deficiency of the experimental design also complicates the overyielding assessment: Hector et al. failed to account for the eightfold increase in the planting density of each species from the highest to the lowest species richness treatments. This flaw, an inevitable consequence of the “substitutive” design of the experiments, introduces a hidden treatment into the experiment that results from intensified intraspecific competition at higher planting densities (6–11), which correspond to lower species richness. The gradient in planting densities (from 250 to 2000 seeds m\(^{-2}\)) of individual species across the diversity treatments invalidates the attribution of any differences in individual plant size [table 4 of (1)] or total biomass per species to the effects of the species richness treatments. An experimental design that included both additive and substitutive planting designs (7, 10) could potentially overcome the confounding factors of inadequate replication and lack of control for planting density.

Some of the uncertainty that results from design deficiencies in these experiments could have been reduced by a more complete reporting and analysis of both the treatments and the responses. Reporting the individual species biomasses in the mixtures and using statistics that represent relative abundance would have quantified the actual treatment levels more accurately than the use of “planned species richness,” which overestimated the number of species that actually grew in each plot [note 19 of (1)] and assumed that all species contributed significantly to the measured response (3–5). Similarly, use of the observed numbers of individuals alive when the data were collected, rather than the number of individuals sown in year 1, would have allowed valid conclusions to be drawn about both individual plant size [table 4 of (1)] and the mechanisms that produced differences in species responses across the experiment. Information on the species composition of all mixtures would have permitted identification of mixtures with true overyielding and facilitated evaluation of the mechanisms that produced the overyielding.

All of these issues, as well as several others, have been thoroughly discussed in published comments on earlier experiments of this type (4, 5, 7, 13, 14, 17–19). We agree with Hector et al. that environmental conditions have a major effect on plant productivity and that overyielding does occur in some multispecies mixtures, particularly those containing nitrogen-fixing legumes. In light of our analysis of the data presented in their figure 2, however, we conclude that species richness per se has no statistically or biologically significant effect on plant productivity across the eight sites of the BIODEPTH experiment.

M. A. Huston
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831–6335, USA
E-mail: hustonna@ornl.gov

L. W. Aarsn
Department of Biology
Queens University
Kingston, ON K7L 3N6, Canada

M. P. Austin
Division of Wildlife and Ecology
Commonwealth Scientific and Industrial Research Organisation
Post Office Box 84
Lyneham, ACT 2602, Australia

B. S. Cade
Biological Research Division
U.S. Geological Survey
4512 McMurry Avenue
Fort Collins, CO 80525–3400, USA
level of species richness with the biomass produced by the most productive monocultures. To estimate the maximum productivity found in monocultures at each site, we used both replicates of the species that produced the maximum biomass of all monoculture replicates (the experimental design included only two replicates of each species in monoculture). At five of the eight sites, the two replicates with highest biomass belonged to a single species; at the remaining three sites, the other replicate of the species with the highest biomass ranked lower than second (third, third, and sixth, respectively). We included all replicates in the mixtures with the highest appropriate species richness (i.e., the species richness treatment that was half the number of monoculture plots, based on table 4 of (1)), because all replicates at this level of richness should have been identical in species composition. For the species richness treatments between one and the maximum described above, we used combinatorial probabilities to determine what proportion of the plots at each level were likely to include the most productive species from the monocultures (no information on the actual species composition of the mixtures was released with the article). Although only three of the eight sites were designed properly for analysis of overyielding, we performed this analysis for the other five sites as well. At each treatment level of species richness, we sorted the communities from highest to lowest biomass, and selected the number of replicates corresponding to the combinatorial probability described above, beginning with the highest biomass. Thus, our approach was designed to use the maximum number of appropriate replicates at each treatment level, and was biased toward detecting overyielding by selecting the most productive mixtures. If increasing species richness causes higher productivity through overyielding, the slope of the regression through these points should be positive. Alternative statistical approaches, such as quantile regression [B. S. Cade, J. T. Terrell, J. Scheller, Ecology 80, 313 (1999)], can be used to quantify the maximum response of productivity to increasing species richness. Analysis of the most productive plots in each treatment at each site (90th percentile) indicates that the 90% confidence intervals for regression slope overlap among all the sites (B. S. Cade, unpublished analysis).

Response: The goal of BIODEPTH (1) was a general test of the relationship between biodiversity and ecosystem functioning in a range of European grasslands. Therefore, we conducted standardized experiments replicated at multiple locations, using combined analyses to test directly for differences between sites [figure 1 of (1)]. Huston et al. question the lack of statistically significant differences in species effects between sites in our combined analysis. Had those differences been strong enough, however, our combined analysis would have shown a significant location-by-richness interaction with different regression slopes. The interaction instead was not significant [p. 1124 of (1)], and the mean square for the significant species richness effect was 30 times larger, which supported emphasis on the main effect (2). Our richness tests take remaining compositional variation between communities with the same diversity into account (3), because richness is tested against the assemblage term [table 3 and note 22 of (1)]. Effects of species richness were subdivided into a significant log-linear term and nonsignificant deviation, which provided objective grounds for consistency (4), despite differences between individual-site analyses.

Huston et al. also note that mixtures could have been dominated by highly productive species not grown in monoculture. Only 37 of the 308 polycultures were dominated by nonmonoculture species [note 33 of (1)], however. Excluding plots dominated by nonmonoculture species does not alter conclusions in Switzerland (5) or across all sites; analyses with only mixtures composed entirely of monoculture species likewise produced a significant log-linear species richness effect (F1,24 = 9.21; P < 0.001). More generally, there is no reason that the requirement to have all possible assemblages should be restricted to monocultures. Doing so arguably would bias the experiment, because just as a species with a higher yield than any planted monocultures could be missed, so too could mixtures that would be more productive than chosen polycultures. Growing the myriad combinations of a diverse set of species is not feasible, so random designs select a representative subset. Excluding monocultures from analyses as suggested misses important points (6); still, after limiting the diversity gradients, individual site biomass regressions reveal positive productivity–diversity relationships (solid lines in figure 1 of Huston et al.). Overyielding analyses (broken lines in figure 1 of Huston et al.), however, sometimes produce negative regressions, which indicate progressively stronger underyielding. Such underyielding, which could be produced by increasing interference or allelopathy, contradicts the pattern of increasing productivity. Mismatch of pattern and suggested mechanism arises in this case, however, because the alternative analysis suggested by Huston et al. (their note 16) uses several polycultures at each level of diversity but only the single highest-yielding monoculture; consequently, mixtures are often compared with monocultures of highest-yielding species that they do not include, confounding changes in richness and composition (7). Although Huston et al. argue that the pattern of increasing productivity with increasing species number could stem from the effects of random sampling, standard relative yield total (8, 9) and related techniques (10) reject the sampling effect and are consistent with complementary and positive interactions [sampling effect null prediction: RYT = 1; mean over all polycultures = 1.5, SEM = 0.06, t test = 25.3, P < 0.001, n = 204 (11)].

The proposed mechanisms underlying diversity effects are the functional traits of individual species and groups. We reported, as Huston et al. reiterate, that although the contributions of most individual species to the effects of biodiversity were small, those of Trifolium pratense were large (1). We expected complementary and positive effects of nitrogen-fixers to play a large role in the species richness effect; hence, legumes were an a priori functional group in our design (1, 5). There is no magic effect of “species richness per se.” In principle, all effects can be attributed to the traits of individual species and their interactions. There is no reason that there cannot be effects of both richness and composition; indeed, if there were no effects of composition (i.e., if individual species effects on functioning were identical), there could be no effect of species richness.

Notwithstanding the suggestion to the contrary by Huston et al., our regressions of estimated average plant size (I) did account for decreasing density of individual species with increasing species number. Because individuals of clonal species are hard to quantify, we divided the biomass of a species in a mixture by the number of seeds sown to gain estimates of individual performance. Dead individuals entered the calculation of average size as zeros. Regression slopes show how the size of an individual plant changes as more of its neighbors belong to other species.

Intense intraspecific competition in monocultures and low-diversity mixtures is not a “hidden treatment,” as Huston et al. suggest, but is one potential biodiversity mechanism (1, 12): benefits of biodiversity arise if, because of complementarity, interspecific competition is less intense on average than intraspecific competition. Substitutive experimental designs take monocultures and introduce biodiversity by replacing conspecifics with other species. Both substitutive and additive designs have their strengths and limitations (12, 13). Classical additive and substitutive experiments often have restrictive physical situations (15) and short time scales that can restrict complementary interactions. Our experiment was conducted under field conditions, with more natural communities, over several years, which allowed plants to grow and reproduce. Results of substitutive experiments are most often insensitive to changes in total density (13), and we doubt that ours are specific to the initial total sowing density; a companion experiment varying total density of a three-species mixture found no effect of initial density on community biomass by the second year after sowing (16).
Huston et al. argue that we did not adequately report on relative species abundances, and object to our use of measures such as sown species richness. However, our tests of the sampling effect and complementarity took relative abundances of species into account (9, 10, 11). Experimental responses can result from processes that integrate over time and space, with initial and realized abundances over whole plots and entire time courses often contributing (8). The close match of sown and realized diversity supports analyses using sown species levels; using observed numbers produces similar results [note 19 of (1)]. Diversity indices likewise yield similar results: the Shannon index and its exponent, which gives the effective number of equally abundant species, both have highly significant positive relationships with productivity.

Finally, we take this opportunity to clarify an issue raised only in passing by Huston et al.: whether our experimental approach of random species addition or removal accurately reflects processes in natural or human-influenced systems. Extinction often is not random, and depends on the process driving it, and is inconsistent with other suggested designs (18, 19). The research they cite (their references 3 through 5) refers to particular scenarios only, and we suspect that good general evidence is currently lacking on orders of loss in relation to a broader set of extinction drivers. For example, recent studies of fragmented grassland remnants (20, 21) have found that ability to produce high yields or to dominate communities did not appear to influence extinction, whereas life cycle characteristics such as seed dormancy and lifespan did. Species loss could be variable and unpredictable, given forecasted high levels of global change and climate alteration, and we clearly need better ecological predictions.

In view of the current, limited evidence, random loss scenarios may provide adequate models in some cases and should be properly tested, not dismissed on assumptions alone. Although, as we acknowledged in our report, our biomass patterns were partly generated amid the variety of special cases. Differences and interactions among species that have functional consequences provide the underlying mechanisms. Variation ultimately forms a continuum, and different categorizations are alternative ways of viewing the same functional biodiversity. Species richness provides one easily measured axis of differentiation, because all else being equal, communities with more species may generally have a greater range of functional differences than depauperate versions of the same community (38).

In nature, species selection is not made by scientists, and there is no guarantee that the species needed to maintain ecosystem processes will remain. Knowing that the combination of two particular species accounts for most of the high yield of a diverse mixture is of little use if one or both have gone extinct. Moreover, conditions vary and change, as does the performance of species, so that the same mixture may no longer perform as well. Although many further experiments are needed, varying species numbers, differences, frequencies, and densities, we must not forget to seek general patterns amid the variety of special cases.

A. Hector
Natural Environmental Research Council (NERC)
Centre for Population Biology
Imperial College at Silwood Park
Ascot, Berkshire GB-SL5 7PY, UK
E-mail: a.hector01@ic.ac.uk

B. Schmid
Institut für Umweltwissenschaften
Universität Zürich
Winterthurerstrasse 190
Zürich CH-8057, Switzerland

C. Beierkuhnlein
Lehrstuhl Biogeographie
Universität Bayreuth
Bayreuth D-95440, Germany

M. C. Caldeira
Departamentos de Engenharia
Florestal e de Botânica
Universidade Tecnica de Lisboa
Tapada da Ajuda
Lisboa PT-1399, Portugal

M. Diemer
Institut für Umweltwissenschaften
Universität Zürich

P. G. Dimitrakopoulos
Biodiversity Conservation Laboratory
Department of Environmental Studies
University of the Aegean
Karadoni 17
Mytilene, Lesbos GR-811 00, Greece

J. A. Finn
Department of Agriculture
University of Reading
Earley Gate
Post Office Box 236
Reading, Berkshire GB-RG6 6AT, UK

H. Freitas
Departamento de Botànica
Universidade de Coimbra
5000 Coimbra, Portugal

P. S. Giller
Department of Zoology and Animal Ecology
University College Cork
Lee Maltings
Prospect Row
Cork, Ireland

P. Högbom
Department of Forest Ecology
Swedish University of Agricultural Sciences
Umeå SE-90183, Sweden

K. Huss-Danell
Crop Science Section
Department of Agricultural Research for Northern Sweden
Box 4097
Swedish University of Agricultural Sciences
Umeå SE-90403, Sweden

J. Joshi
Institut für Umweltwissenschaften
Universität Zürich

A. Junnponnen
Division of Biology
125 Ackert Hall
Kansas State University
Manhattan, KS 66506, USA

C. Körner
Institute of Botany
University of Basel
Schoenheinstrasse 6
Basel CH-4056, Switzerland

P. W. Leadley
Ecologie des Populations et Communautés
Université Paris Sud XI
URA CNRS 2154
Bâtiment 326
Orsay Cedex FR-91405, France

M. Loreau
Laboratoire d’Ecologie
UMR 7625
Ecole Normale Supérieure
46 Rue d’Ulm
F-75230 Paris Cedex 05, France
A. Minns  
NERC Centre for Population Biology  
Imperial College at Silwood Park

C. P. H. Mulder  
Victoria University of Wellington  
School of Biological Sciences  
Post Office Box 600  
Wellington, New Zealand

G. O’Donovan  
Department of Environmental Resource Management  
University College of Dublin  
Belfield  
Dublin, Ireland

S. J. Otway  
NERC Centre for Population Biology  
Imperial College at Silwood Park

J. S. Pereira  
Departamentos de Engenharia Florestal e de Botânica  
Universidade Tecnica de Lisboa

A. Prinz  
Lehrstuhl Biogeographie  
Universität Bayreuth

D. J. Read  
Department of Animal and Plant Sciences  
University of Sheffield  
South Yorkshire GB-S10 2TN, UK

M. Scherer-Lorenzen  
E.-D. Schulze  
Max Planck Institute for Biogeochemistry  
Postfach 10 01 64  
Jena D-07701, Germany

A.-S. D. Siamantziouras  
Biodiversity Conservation Laboratory  
Department of Environmental Studies  
University of the Aegean

E. Spehn  
Institute of Botany  
University of Basel

A. C. Terry  
Department of Environmental Science  
University of Bradford  
West Yorkshire BD7 1DP, UK

Technical Comment

A. Y. Troumbis  
Biodiversity Conservation Laboratory  
Department of Environmental Studies  
University of the Aegean

F. I. Woodward  
Department of Animal and Plant Sciences  
University of Sheffield

S. Yachi  
Centre for Ecological Research  
Kyoto University  
Otsu 252-2113, Japan

J. H. Lawton  
NERC Centre for Population Biology  
Imperial College at Silwood Park

References and Notes

6. Monocultures help identify mechanisms such as reductions in competition in mixtures compared with monocultures due to complementarity. Species not originally sown were weeded because they would not reinvade if extinct in nature. Although some monocultures did require more weeding, all communities required some. Fertilized and naturally productive environments often have low diversities but, as far as we know, none of the species we used form monocultures under conditions similar to those of our experiment.
14. With additive designs, decreasing species number is confounded with decreasing total plant density; each species starts with the same density in all diversity mixtures and does not have to compensate for reduced numbers by producing larger individuals to maintain constant yield. Higher mortality may occur with increased total density (8), which makes it easier for some species to gain dominance. Substitutive designs keep total density constant but are limited to the chosen density.