

# THE MEASUREMENT OF SPECIES DIVERSITY

♦4079

*Robert K. Peet*

Section of Ecology and Systematics, Langmuir Laboratory, Cornell University, Ithaca,  
New York 14850

## INTRODUCTION

Inspired significantly by the provocative papers of MacArthur (34, 35, 38) and Hutchinson (22), ecologists over the past twenty years have devoted considerable energy to the explanation of patterns of diversity in ecologic systems. Despite considerable interest, however, no generally accepted definition of diversity has emerged. "Diversity per se does not exist," was the contention of Hurlbert (20), who suggested abandoning the term because of the multiplicity of meanings and interpretations attached to it. MacArthur (37) also considered the term had outlived its usefulness, and Eberhardt (10), Austin (2), and McIntosh (40) all complained of the lack of a definition. Eberhardt considered that diversity "mostly suggests a considerable confusion of concepts, definitions, models, and measures (or indices)." If diversity is to continue to play a productive role in ecological investigations, agreement is needed on the definitions of the many constituent concepts included in its current application.

At the community level of synthesis many phenomena are complex and open to multiple interpretation. Consequently, many authors have suggested diversity indices appropriate for their own studies, no one of which can be considered a priori correct for general application (12, 19, 20, 32, 40). Diversity, in essence, has always been defined by the indices used to measure it, and this has not fostered the sort of uniformity which allows the clear statement of ideas and hypotheses. Progress in ecology, as in all science, depends upon precise and unambiguous definition of terms and concepts (2, 40, 44). The present contribution attempts to define in a precise, but still generalized manner, what is or should be meant by the many terms surrounding the concept-cluster *diversity*. Guidelines are also suggested for the application of the many available diversity indices.

## ASSUMPTIONS AND DATA

The measurement of ecologic diversity is not as simple as might be expected. A number of assumptions and decisions regarding the data to be analyzed are required.

Measurement of diversity requires a clear and unambiguous classification of the subject matter. In ecology reference is usually made to species diversity, but nothing precludes treatment of varietal, generic, or even structural diversity. Pielou (53) has discussed some of the problems involved in the simultaneous treatment of different levels of a hierarchic classification.

All individuals assigned to a specific class are assumed equal. Preston (59) has pointed out the difficulties in assuming equality of forms (e.g. sexes, larval stages) which play very different functional roles in a community. Hendrickson & Ehrlich (18, also 16, 28) have gone further, suggesting modification of existing indices to account for within-species variation.

All species or classes are assumed to be equally different. Lloyd (30) was the first to question the assumption of species equivalency when he suggested a correction factor to compensate for seasonal variation in diversity due to breeding behavior. Johnson & Raven (23) also questioned the assumption and Hendrickson & Ehrlich (18) proposed that indices be modified to include such variation.

Most diversity indices require an estimate of species importance. The actual measure will depend on the particular question being studied, but the choice can greatly influence the results obtained, as Dickman (9) has amply demonstrated.

Many organisms, particularly those of sessile habit, exhibit pattern in their distribution. This necessitates careful randomization in sampling (12, 40, 61). Pielou (53) has addressed this problem in much greater detail.

Whittaker (68-72) distinguished three levels of diversity. His alpha diversity is the within-habitat or intracommunity diversity that is the subject of the present contribution. Beta or between-habitat diversity is defined as the change in species composition along environmental gradients and can be most easily measured in terms of half-changes. Gamma diversity is the diversity of an entire landscape and can be considered a composite of alpha and beta. These forms are not always easily distinguished. Many alpha diversity measurements are influenced by habitat variation, which could be interpreted equally well as beta diversity.

## CONCEPTUAL APPROACHES AND TERMINOLOGY

Much of the confusion surrounding the measurement of diversity is the direct result of confounding several distinct concepts along with loose application of the now burdensome terminology. To understand the literature of the subject, however, it is first necessary to understand that terminology and to grasp the distinctions between the principal concepts involved.

### *Species Richness*

The oldest and most fundamental concept of diversity is species number. Fisher, Corbet & Williams (13), who were among the first to seek to quantify this concept, employed the parameter alpha in the logarithmic relation of numbers of species to numbers of individuals as an index of diversity. When it became obvious that several concepts were implied by the term diversity, Lloyd & Ghelardi (31) suggested this

aspect be called species number as opposed to a second aspect termed equitability or evenness. As a term *species number* proved unpopular, probably because it implies that the number of species in a community can actually be determined. While the number of species in a sample is easy to ascertain, a natural community is an open system with nothing approaching a fixed number of species. McIntosh (40) suggested the alternative term of *richness* for the concept of species number. Species richness is in frequent use today and is perhaps the least ambiguous of all the diversity terminology.

A number of less important terms have been applied to species richness. Hurlbert (20) designated the number of species occurring per unit area as species density, while Auclair & Goff (1) called such indices variety indices. Species per fixed number of individuals is richness in the strict sense according to Hurlbert, but Auclair & Goff call these species/individuals indices.

### *Heterogeneity*

Should a community with five equally abundant species be considered to have the same diversity as a similar community with the same five species, one of which comprises 95% of the individuals? According to a second concept of diversity which compounds richness with evenness, the answer is no. Indices of this diversity concept measure not the absolute number of species in the community, but the functional or apparent number of species. For example, two individuals selected at random are much more likely to represent different species if taken from the first of the above populations than if taken from the second. From this perspective the community with five equally common species appears to have more species despite the equal species counts.

This dual-concept diversity was introduced into the ecologic literature by Simpson (65), who had become aware of a similar approach in the work of Yule (76) on the statistical analysis of vocabulary. Many contributors have considered that diversity should include both an evenness and a richness component as Simpson had implicitly suggested. This concept has now come to be synonymous with diversity for many workers. While Hurlbert (20) considered that the term diversity should be restricted to this concept if it is to retain any meaning, it seems unlikely that all the other connotations will quickly be forgotten. Retention of diversity as a broad term encompassing all of the subordinate concepts, and specification of the dual-concept or mixed-diversity measures in some other manner seems more desirable.

Whittaker (69, 71) has advocated using the Simpson index to express relative concentration of dominance. Sanders (61) reached a similar conclusion on the dual-concept measures calling them dominance diversity indices. In his 1972 review Whittaker refined his terminology, calling all such indices measures of slope (of the importance value sequence), but differentiating between Simpson's index for concentration of dominance and Shannon's formula as an index of equitability. Auclair & Goff went further, calling all these indices equitability indices. This use of equitability can be confusing; the term is more frequently used to denote a quite different concept, discussed in the next section. The term *heterogeneity*, suggested by Good (15) and Leti (27), appears less ambiguous and will be used throughout this review.

### *Equitability*

Having considered both species richness alone and a heterogeneity concept that combines richness with evenness, it is an obvious extension to try to quantify that component of evenness. Lloyd & Ghelardi (31) were the first to discuss formally such an evenness component, though Patten (49) and Margalef (42) used an index of redundancy to measure a very similar concept. Lloyd & Ghelardi explained that the term *evenness* refers to the absolute evenness of a distribution, while *equitability* refers to evenness relative to any specific standard, such as the broken-stick model of MacArthur (34). Evenness measures can thus be considered a subset of measures of equitability. Sheldon (63), in contrast, has called the equitability concept *relative diversity*, a term Kohn (25) has applied to a form of heterogeneity index.

## SPECIES RICHNESS INDICES

Richness is an indicator of the relative wealth of species in a community. While the concept is simple, it is almost impossible to provide a formal definition. The difficulty stems from the inherent dependence of any richness measure on sample size; the larger the sample the greater the expected number of species. Because it is virtually impossible to ascertain the complete composition of an ecologic community, richness is often measured as the number of species in samples of an arbitrarily chosen constant size. It would clearly be desirable to have a richness index independent of sample size.

### *Simple Indices*

Traditionally a series of simple indices have been employed as measures of richness independent of sample size. All such indices presuppose a particular functional relationship between the expected number of species observed and the sample size. Certain parameters of these functions act as richness indices. For example, if we assume that the expected number of species  $E[S]$  is equal to a constant  $k$  times the square root of the number of individuals  $N$  in the sample, where  $k$  varies between communities, then  $\hat{k} = S(N)^{1/2}$  is an estimate of  $k$ , the richness of the sample. In general, if we are studying a set of communities (or sample universes) with a known functional relationship between the expected species number and the sample size [i.e.  $E[S] = f(k, N)$  where  $k$  is the unknown parameter indexing richness], then we can find an estimate  $\hat{k}$  for that richness. Such a measure escapes commitment to an arbitrary sample size by expressing richness as the rate at which the number of species or sample size increases.

Two assumptions are implicit in the use of such an index. First, the functional relationship between the expected number of species  $E[S]$  and the number of individuals in the sample  $N$  remains constant among the communities being studied. Second, the precise functional relationship is known. If these assumptions are not satisfied, the index of richness will vary as a function of sample size in some unpredictable manner (17, 20, 55, 71).

Numerous species-individuals relationships have been utilized as bases for richness indices. Margalef (41, 42) suggested a logarithmic relationship:  $R_1 =$

$(S - 1)/\text{Log } N$ . A similar index was proposed with different motivation by Odum et al (48):  $R_2 = S/\text{Log } N$ . Menhinick's (45) square root relation has already been mentioned:  $R_3 = S/(N)^{1/2}$ .

Two of the most frequently used and thoroughly investigated models of the relationship between species richness and the relative abundance of species (73) are the log series of Fisher, Corbet & Williams (13) and the lognormal distribution of Preston (56). Fisher et al suggested that the numbers of individuals representing different species follow a logarithmic series. Summing over the series they obtained the relation  $S = \alpha \text{Log}(1 + N/\alpha)$ , where  $\alpha$  is a fitted constant indexing diversity.

Preston's proposal (56-58) that a lognormal distribution provides the best fit for species abundance data is well known, and subsequent work has shown the distribution to fit many different kinds of community samples (73). Preston (56-58) used the parameters of this relationship to calculate the expected number of species in a total sample or universe. His expression was  $S^* = Y_o \tilde{\sigma} (2\pi)^{1/2}$ , where the  $Y_o$  is the number of species in the modal octave and  $\tilde{\sigma}$  is the logarithmic standard deviation.  $S^*$  can clearly be used as an index of richness (11, 36).

The dominance diversity curves proposed by Whittaker (69-71) provide a useful perspective for examining diversity patterns (cf Figure 4). These graphs are constructed such that the ordinate represents the logarithm of some importance value (e.g. abundance), while the abscissa is simply the ordered species sequence from most to least important. Whittaker has given careful consideration to the shapes of these curves and their implications. One aspect he considers particularly important is the average slope, for which he has proposed two indices (71). The first is the total number of species encountered,  $S$ , divided by the difference between the logarithmic importance of the most common  $S_1$  and rarest  $S_N$  species. This is a measure of the average number of species per log-cycle of importance, or simply the slope of a straight line connecting the most and least abundant species:  $Ec = S/(\text{Log } S_1 - \text{Log } S_N)$ . The second index, a refinement of the first, is the total number of species divided by four times the logarithmic standard deviation of the importance value:  $Ec' = S/4 [\sum_{i=1}^S (\text{Log } p_i - \text{Log } \bar{p})^2 / S]^{1/2}$  where  $p_i$  is the percentage importance of the  $i^{\text{th}}$  species and  $\bar{p}$  is the geometric mean. This second index, which assumes a lognormal relationship, should theoretically eliminate much of the error in  $Ec$  resulting from use of extreme values of importance, for it divides the species number by the range which would be expected to be covered by 95% of the species.

Species per log-cycle, which the above two slope indices measure, provides a different approach to species richness. While the simple richness indices were based only upon species number and sample size, the slope indices incorporate the species abundances as well. However, because the shape of the dominance diversity curve is influenced by the underlying niche division pattern, these two indices are subject to the same limitations the other simple richness indices are.

### *Species Counts*

The two assumptions necessary for the use of indices relating species number to sample size are rarely satisfied. An alternate approach is provided by direct counts of species numbers in samples. MacArthur (36), Poole (55), and Williamson (74) considered these to be among the most effective richness measures, and Whittaker

and Woodwell (71, 72, 75) found the average number of species per sample to be the best index for the forests they were studying. Direct species counts, while lacking theoretical elegance, provide one of the simplest, most practical, and most objective measures of species richness.

Comparison of species counts requires equal sample sizes. One method of avoiding incompatibility of measurements resulting from samples of different sizes, the rarefaction method of Sanders (61), is to calculate the number of species expected from each sample if all the samples were reduced to a standard size (such as 1000 individuals). While Sander's rarefaction approach is very useful, the original formulation was mathematically imperfect (12, 20, 64). The correct, unbiased form of the expected number of species in a sample size  $n$  drawn from a population of size  $N$  which had  $S$  species was given by Hurlbert (20) as

$$E[s] = \sum_{i=1}^S \left\{ 1 - \left[ \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right] \right\}$$

where  $N_i$  represents the number of individuals in the  $i^{\text{th}}$  species in the full sample

Unfortunately, the species count does not provide a useful method for making inferences about the underlying community. Two communities can possess very different relationships of species importances, yet have the same numbers of species in samples of a particular size. Consider sampling from two communities with 1000 individuals each, the first with three species divided 1:1:1, the second with eleven species divided 90:1:1:1:1:1:1:1:1:1:1. Figure 1 illustrates the relationship between the expected number of species and the sample size. While the second universe is obviously the richer in species, on the average the first population will appear richer for sample sizes less than about 23 individuals. When using a species count method, one assumes that the communities under consideration do not differ too widely in their species-individuals relationship, though one need not worry about the actual functional relationship involved. In conclusion there is no real basis for comparing the richness of a series of communities using a single index unless one can assume that the underlying species-individuals relationships are similar.

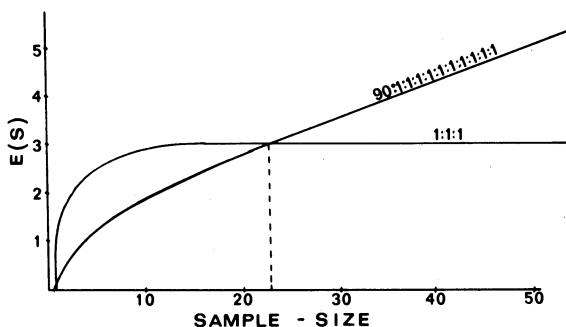


Figure 1 The relationship between the expected number of species and sample size for samples drawn from two hypothetical communities of 1000 individuals.

## HETEROGENEITY INDICES

Margalef and Pielou, two of the strongest advocates of the heterogeneity (dual-concept diversity) approach have both stated how they consider diversity should be defined. Margalef (43) explained that "... diversity is a statistical function that implies no particular regularity in distribution, and in whose computation the numbers of individuals in all the species are taken into account." Pielou (54) wrote "... diversity, however defined, is a single statistic in which the number of species and the evenness are confounded." One point appears agreed upon: there should be two contributing components, the number of species and the distribution of individuals among those species (equitability).

### *Simpson's Index*

The first of the heterogeneity indices used in ecology was proposed by Simpson (65). His index measures the probability that two individuals selected at random from a sample will belong to the same species. For an infinite sample this index is

$$\lambda = \sum_{i=1}^S p_i^2$$

or for the finite sample case,  $L = \sum [n_i (n_i - 1)] / [N (N - 1)]$  where  $p_i$  is the proportion of the individuals in species  $i$ ,  $n_i$  the number of individuals in species  $i$ , and  $N$  the total sample size. This index can alternatively be interpreted as the weighted mean of the proportional abundances, as Hill (19) and Leti (27) have suggested.

Simpson's index as originally formulated varies inversely with heterogeneity. To avoid this difficulty Greenberg (16) and Berger & Parker (3) proposed subtracting the Simpson index from its maximum possible value of 1, a formulation originally suggested by Gini (14):  $D = 1 - \sum p_i^2$ . Pielou (54) has suggested that it is statistically more correct to use a formulation adjusted for finite sample size:

$$\tilde{D} = 1 - \sum \{ [n_i (n_i - 1)] / [N (N - 1)] \}$$

Hurlbert has also suggested this index but in a somewhat different form under the name of "the probability of interspecific encounter." A geometric interpretation of diversity as measured by indices of this form has been proposed by Bhargava & Doyle (4).

Williams (73) suggested using the reciprocal of Simpson's index and MacArthur (37) suggested that for certain theoretical models the reciprocal has particularly desirable mathematical qualities. In this form the index can be interpreted as the number of equally common species required to produce the same heterogeneity as observed in the sample (19, 20, 39). This index has also been used in population genetics as a measure of the effective number of alleles in a population (8).

McIntosh (40) has proposed an interesting variation of Simpson's index. A community can be represented as a point in an  $n$ -dimensional hyperspace where each dimension refers to the abundance of a particular species. In such a space the distance between a community and the origin can be measured using the Pythagorean theorem as  $(\sum n_i^2)^{1/2}$ . The greater the number of individuals in a particular

species, the farther the stand will be from the origin; and if the individuals are spread out more evenly between species or their number reduced, the distance will also be reduced. That is to say, the more heterogeneous a stand, the closer it will be to the origin. McIntosh suggested that a useful measure of heterogeneity would be one that related the distance between a stand and the origin to the range of possible values as determined by the number of individuals in the sample. The resultant index is

$$D_{mc} = (N - \sqrt{\sum n_i^2}) / (N - \sqrt{N})$$

where  $N$  is the total number of individuals and  $n_i$  is the number in species  $i$ . When percentage importance is used in preference to numbers of individuals, McIntosh's index simplifies to  $1 - (\sum P_i^2)^{1/2}$ . This is merely a square root transformation of Gini's index. Bullock's (6) harsh criticism of McIntosh's index appears to be unfounded, since he based it on  $(\sum n_i^2)^{1/2}$  rather than the standardized form that McIntosh recommended.

### *Information Theory Indices*

The most popular of the heterogeneity indices are those based on information theory. This approach was first applied to ecology by Margalef (42) and has since gained much support. Pielou (51) explained the application of information theory to diversity measurement, suggesting that heterogeneity can be "... equated with the amount of uncertainty that exists regarding the species of an individual selected at random from a population. The more species there are and the more nearly even their distribution, the greater the diversity." The argument concludes that since information content is a measure of uncertainty, it is a reasonable measurement of heterogeneity.

The expression for the information content per individual within an infinite population is given by the Shannon-Weaver formulation (62):

$$H' = - \sum_{i=1}^S p_i \text{Log } p_i$$

where  $p_i$  is the percentage importance. Expression of heterogeneity in terms of the antilogarithm of  $H'$  simplifies interpretation.  $\text{Exp } (H')$  measures the number of equally common species which would produce the same heterogeneity or  $H'$  as the sample (7, 19, 36). Alternatively,  $1/\text{Exp } (H')$  can be interpreted as the geometric mean of the proportional abundances (19, 27). Whittaker (71) has suggested additional reasons for using the antilog form.

In actual application  $H'$  is frequently estimated using

$$h' = -\sum (n_i/N) \text{Log } (n_i/N)$$

The use of  $n_i/N$  as an estimate of  $p_i$  results in a biased estimate. Bowman et al (5, also 21) have calculated the expected value of  $h'$  as

$$E[h'] = -\sum p_i \text{Log}_e p_i - \frac{S-1}{2N} + \frac{1 - \sum p_i^{-1}}{12N^2} + \frac{\sum (p_i^{-1} - p_i^{-2})}{12N^3} + \dots$$

Fortunately the bias is small for most ecological applications. An additional error in  $H'$  is introduced if all species in the community are not included in the sample.



In most ecological situations this error will also be small, but a large sample size helps insure this.

Bowman et al (5, also 21) have also provided a formula for the variance of  $h'$ . This can be used in a test for equality of  $H'$  from two samples following a method proposed by Hutcheson (21). However, the fact that the calculations for variance also require knowledge of the number of species in the sampling universe greatly reduces the applicability of this procedure to ecologic situations.

If heterogeneity is equated with uncertainty, the Shannon-Weaver formulation is a biased indicator valid only for an infinite sample (50-53). The correct formulation for the finite sample is given by the Brillouin formula  $H = (1/N) (N! / \prod_{i=1}^S n_i!)$ , from which the Shannon formula can be derived through the use of Stirling's approximation for  $\log N!$ . Lloyd et al (32) have provided a calculation formula for this index. Pielou (54) gives clear derivations of both  $H$  and  $H'$ .

Pielou (53) argued for use of the Brillouin formula in preference to the Shannon formula on the grounds that the latter does not reflect sample size; the same result is always obtained if the species proportions are kept constant. She claimed that an index should vary as a function of the sample size as does the Brillouin index. An example from Peet (manuscript) illustrates the response of the Brillouin index to changes in sample size. Consider two hypothetical communities: the first composed of ten species, each represented by five individuals; the second with nine species, eight represented by 110 and one by 120 individuals. Using the Brillouin formula 0.874 is obtained for the first community and 0.943 for the second. This implies that the second is the more diverse, yet the first community appears more diverse. It has both the greater species richness, with ten species and fifty individuals compared with nine species and 1000 individuals, and the greater equitability with a perfectly even distribution. According to the evenness and richness criteria the first community is the more diverse, which the Shannon index suggests with values of 1.000 and 0.954 for the two communities respectively. As Pielou (54) also indicated, there is no agreement that uncertainty should be equated with diversity. The above result, showing conflicting interpretations based on the Brillouin and Shannon indices, suggests that uncertainty as measured by the Brillouin formula does not provide an acceptable index of heterogeneity.

### *Theoretical Considerations*

**PROBLEMS IN INDEX SELECTION** The heterogeneity indices thus far discussed exhibit considerable variation in response to selected changes in sample composition. In addition, many other potential indices are available for consideration (4, 15, 19, 27). As an illustration of the number of indices available, consider briefly the sets proposed by Good (15) and Hill (19). Good suggested that heterogeneity can be measured by an index of the form

$$C_{m,n} = \sum_{i=1}^S p_i^m (-\log p_i)^n; (m, n = 0, 1, 2 \dots)$$

For the present purposes there is little reason not to consider fractional values of  $m$  and  $n$  equally valid. Examination of this general equation reveals that Simpson's

index ( $C_{2,0}$ ), Shannon's formula ( $C_{1,1}$ ), and the number of species in the sample ( $C_{0,0}$ ) are included as special cases. Hill proposed a different series of heterogeneity indices of the form

$$N_a = (\sum_i p_i^a)^{1/(1-a)}$$

Here again species number ( $N_0 = S$ ), Simpson's index ( $N_2 = 1/\lambda$ ), and the Shannon formula ( $N_1 = \text{Exp}[H']$ ) can be shown to be special cases.

**RESPONSE THEORY** With these many competing potential indices, a theory of index response is needed to aid in selection. Such a theory should not only verify that indices respond according to stated requirements, but should also delimit differences in their responses to changes in sample composition.

Let  $p_i$  represent the percentage importance of species  $i$ . Any change in  $p_i$  must have a corresponding, equal, and opposite change in  $p_j$  such that the importance values still add to 1.0. Any more complex differences between communities can be considered to result from a series of such simple, opposite, and equal shifts. For example

$$|\Delta p_i| = |\Delta p_j + \Delta p_k|$$

would represent the case where one species increased in importance with a corresponding decrease in importance of two other species. The addition of a new species can be considered as a shift away from  $p_i = 0$ . Therefore, in considering differences in heterogeneity of samples, it is only necessary to consider the effect on an index of two equal and opposite changes in importance.

By definition any shift of two species toward more nearly equal importance is a shift toward greater evenness, while a divergence of two species in importance results in a shift toward decreased evenness. Thus, for the special case of two species starting with equal importance, paired opposite changes in importance should decrease evenness. Qualitatively, this says it is necessary that the combined index contribution of any two species diverging in importance should be less than their original combined contribution. A convex response curve allows this to occur as illustrated in Figure 2. Thus we require that the second derivative of that component of the heterogeneity index summed over  $S$  be negative over the entire range of 0–100% importance. It is important to note that if an inflection point were to occur anywhere between 0–100%, it would be possible for a pair of diverging species to cause an increase in heterogeneity.

A class of dominance indices can also be defined. For these indices divergence of equally important species consistently results in an increase in the index value. In such a situation the second derivative of the response curve function will always be positive (Figure 3). The inverse and complement of such indices can frequently be considered as heterogeneity indices.

For both the dominance and the heterogeneity indices it is important that those rare species not actually sampled should not contribute to the index response, otherwise the index would be incalculable because of the impossibility of quantifying species not sampled. Equivalently, the response curves of both dominance and heterogeneity indices must start at the origin.

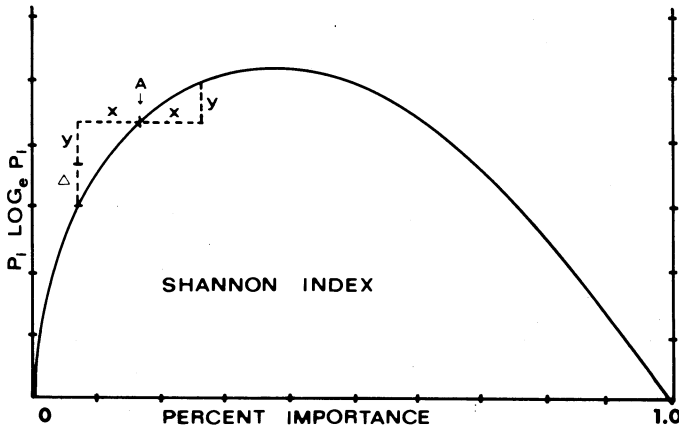


Figure 2 The response curve of a heterogeneity index, represented by the Shannon formula. A divergence of two equally common species from any point *a* results in a net loss ( $\Delta$ ) in heterogeneity.

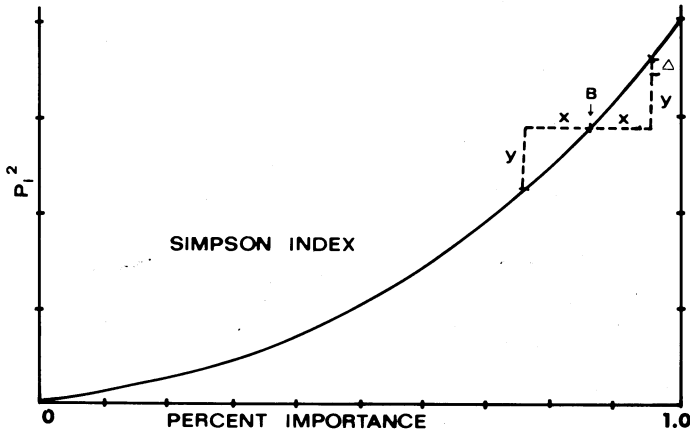


Figure 3 The response curve of a dominance index represented by the Simpson index. A divergence of two species from any point *b* results in a net gain ( $\Delta$ ) in dominance.

Using this theory any number of indices of heterogeneity could be derived. For example, an index might be based on a sine function response curve [ $D_{\sin} = \sum_{i=1}^S \sin(\pi \cdot p_i)$ ] or a response curve with the form of a semicircle [ $D_{\text{cir}} = \sum_{i=1}^S [p_i(1-p_i)]^{1/2}$ ]. A simple parabola through 0 and 1 provides an index identical to Gini's index or the complement of Simpson's index:  $D_{\text{par}} = \sum_{i=1}^S (p_i - p_i^2) = 1 - \sum p_i^2$ . These three indices and many more are sensitive to changes in both sample richness and equitability. Index selection, therefore, requires a more careful examination of the response curves.

**INDEX COMPARISONS** One of the most significant characteristics of a potential heterogeneity index is its response to changes in species importance. For example, it is useful to be able to distinguish whether an index will be most sensitive to changes in the rare or the common species. One of the most effective methods of examining this is to compare changes in index values that result from a set divergence in importance of two equally common species. A graphical solution to this can be obtained using the methods illustrated in Figures 2 and 3. Alternatively, this behavior can be examined numerically by a judicious choice of examples. The response can also be examined analytically using a Taylor series expansion of the response curve functions. The sum of the second and third terms of the series can be used to approximate the change in value of the function resulting from a small change  $d$  in  $p_i$ . Comparison of two opposite but equal changes in  $p_i$  for small but finite values of  $d$  can indicate the effective changes in heterogeneity.

An additional indicator of index response can be found in the value of the second derivative near zero importance. If the absolute value of the second derivative is constant or decreases with decreasing  $p_i$ , the index will become less sensitive to changes in importance as rarer species are considered. In contrast, if the absolute value of the second derivative increases with decreasing  $p_i$ , the index will become increasingly sensitive to changes in importance as rarer species are examined.

The preceding methods can be used to define two distinct types of heterogeneity indices. Those indices most sensitive to changes in the rarest species and characterized by second derivatives increasing as importance approaches zero will be called type I indices. Typical examples include the Shannon index illustrated in Figure 2 and the semicircle index. Type II indices can be distinguished as being most sensitive to changes in the importance of the most abundant species and by having constant or decreasing second derivatives for values of  $p_i$  approaching zero. The Gini, McIntosh, and sine indices are examples of this type. Conceivably, a third class of indices could be defined composed of indices most sensitive to changes in species of intermediate values, but none of the indices in use at present are of this type. While heterogeneity indices do form an intergrading series or continuum with respect to sensitivity to changes in the rare and common species, the behavior of the second derivative near zero importance provides a convenient and conceptually useful criterion for separation.

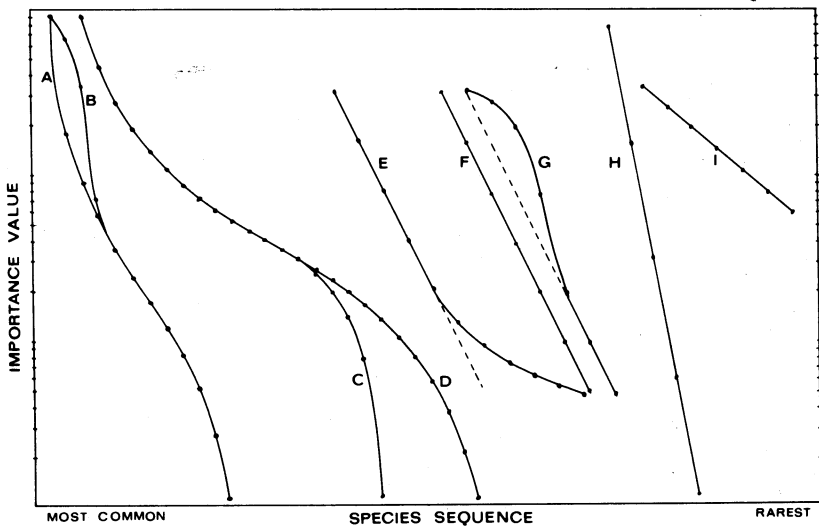
The characteristics of index response have not always been appreciated. Monk (46) and Sager & Hasler (60) have criticized the Shannon index for being insensitive to rare species. The preceding discussion of response behavior suggests the opposite interpretation. Fager (12), Whittaker (71), and Poole (55) suggest that the Shannon index is most sensitive to species of intermediate importance, or more specifically to those with a percentage importance approaching  $1/e$ . While it is true that the maximum contribution an individual species can make to index response occurs when its importance is equal to  $1/e$ , the index responds most strongly to changes in the importance of the rarest species. This is not to say that elimination of a species with a percentage importance of 0.01 will have an effect greater than, say, a 20% reduction in a species with a 0.5 importance. However, the effect of a change of 0.01 importance for a pair of species with initial importances of 0.01 and 0.5 will be greater for the rarer species.

The response theory also provides an explanation for the irregularities described by Hurlbert (20), who reported opposite trends in index response for two "diversity" indices,  $-\sum p_i \log p_i$  and  $1/\sum p_i^2$ , when confronted with the two very different communities listed in Table 1. The explanation is clearly that the first index is a type I heterogeneity index while the second is of type II.

**Table 1** A comparison of heterogeneity index values from two hypothetical communities (from Hurlbert, 20)

Community	Abundance		H'	$1/\sum p_i^2$
AA	$N = 18,000$	$i = 1, 2$	0.78	5.98
	$N = 16,000$	$i = 3-6$		
BB	$N = 40,800$	$i = 1$	2.70	5.00
	$N = 667$	$i = 2-91$		

Two additional examples will clarify the distinctions in index response. Consider two hypothetical communities of the types illustrated by dominance diversity curves A and B in Figure 4. Species number is kept constant and the rarest species have identical importances; only the degree of dominance by the most abundant species is varied. Table 2 lists the diversity values for this and the following example. As the theory predicts, there is a much greater response to this change by the type II index than by the type I. Curves C and D in Figure 4 illustrate the appropriate counterexample with the dominance being kept constant and the richness being increased through an increase in the number of rare species. Here the trend is reversed with the type I index showing the greater response.



**Figure 4** A series of dominance diversity curves representing hypothetical communities used in the text for illustrating index response.

Table 2 Diversity values for hypothetical communities illustrated in Figure 4

Community	Species <sup>a</sup>	$1/\Sigma p_i^2$ <sup>b</sup>	$\text{Exp}(H')$ <sup>c</sup>	$R_{1:2}$ <sup>d</sup>
A High Dominance	12	2.004	3.177	1.585
B Low Dominance	12	3.190	3.899	1.222
(Percent Change)		(59.20)	(27.70)	
C Low Richness	19	5.530	8.650	1.564
D High Richness	25	5.780	9.506	1.645
(Percent Change)		(4.53)	(9.90)	
E High Richness	11	3.248	4.616	1.421
F Geometric Series	7	2.953	3.830	1.297
G Low Dominance	7	3.643	4.141	1.145
H Geometric Series	7	1.518	1.920	1.265
I Geometric Series	7	5.256	5.951	1.132

<sup>a</sup>Species Count, Hill's  $N_0$ .<sup>b</sup>Reciprocal of Simpson's Index, Hill's  $N_2$ .<sup>c</sup>Exponential of Shannon's Index, Hill's  $N_1$ .<sup>d</sup> $\text{Exp}(H')$  ( $\Sigma p_i^2$ ), Hill's  $R_{1:2}$ .

While the preceding discussion contributes to a better understanding of index response, it does not resolve the problem of index selection. In the end this decision must rest with the individual investigator. In all cases, however, the use of indices that are both simple to apply and ecologically easy to interpret is to be encouraged. The series of indices (or diversity numbers) proposed by Hill (19) meets both these requirements. Hill's numbers measure the functional or apparent number of species in a sample and consequently have species as units rather than bits, probabilities, or other units of questionable ecological meaning. This series also includes both type I ( $N_1 = \text{Exp}[H']$ ) and type II ( $N_2 = 1/\lambda$ ) heterogeneity indices. In most cases calculation of  $N_1$  and  $N_2$  will suffice to answer any question that a heterogeneity index can answer.

## EQUITABILITY INDICES

Incorporated within the heterogeneity concept is a component concept concerning the evenness with which importance is divided among species. This component, termed equitability, is logically independent of the second component concept, species richness. Just as it is frequently of interest to study the species richness of a series of samples, it is also of interest to quantify the evenness component.

### *Simple Indices*

Numerous approaches to the measurement of equitability have been proposed. Among these the most frequent approach has been to scale a heterogeneity measure relative to its maximum possible value when the sample size and species number are fixed. More specifically, two formulations which converge for large samples are

frequently used:  $\text{Evenness} = (D - D_{\min}) / (D_{\max} - D_{\min})$  and  $\text{Evenness} = D / D_{\max}$ , where  $D$  is a heterogeneity value for the sampled population and  $D_{\min}$  and  $D_{\max}$  are the minimum and maximum values possible for the given species number and sample size. Hurlbert (20) has referred to these as  $V$  and  $V'$  respectively. Of these the most well known and frequently utilized are Pielou's  $J$  and  $J'$  (51, 53):  $J = H / H_{\max}$  ( $= V'H$ ),  $J' = H' / H'_{\max}$  ( $= V'H'$ ). Also well known is the redundancy measure applied by Margalef (42) and Patten (49) that uses a somewhat different variation of the same approach, usually in the context of information theory:  $\text{Redundancy} = (D_{\max} - D) / (D_{\max} - D_{\min})$ .

As should be evident from earlier discussion, the response behavior of the heterogeneity indices upon which these equitability measures are based will determine their behavior. An index based on the Shannon formula will give more weight to the equitability of the rarer species than an index based on the Simpson measure.

Recently a number of alternative approaches to measuring equitability have been proposed. Fager (12) suggested two indices; one based on the standard deviation of the number of individuals in a species, the other on the number of moves of single individuals from one species to another which would be required to produce a completely even distribution. Peet (in 71) has suggested an index based on the sum of squares of deviation from an even distribution; an index which can be shown to be equivalent to  $V$  Simpson. Buzas & Gibson (7) observed that  $\text{Exp}(H')$  is an index of the number of species which when perfectly even in distribution will produce the same value of  $H'$  as the sample being measured. They suggest that an appropriate evenness measure would be provided by dividing this number by the actual number of species present. Lloyd & Ghelardi (31) used a similar approach but related the number of species distributed by MacArthur's broken-stick models, which would produce the same value of  $H'$ , to the number of species actually present.

Equitability can also be measured in terms of the underlying species-abundance relation when this is known. Drawing upon Longuet-Higgin's (33) observation that the  $H'$  of a lognormal distribution takes the form  $\text{Log } S - \frac{1}{2} \sigma^2$  ( $S$  = species in universe,  $\sigma$  = logarithmic standard deviation), Edden (11) suggested that  $H'$  values can be split into two components when the lognormal is known to occur. While Preston (58) supposed  $\sigma$  to be relatively constant, it is an obvious measure of population dispersion or equitability. Where applicable, this approach provides one of the most intuitively appealing measures of equitability.

While many of the above equitability indices are widely used, their limitations are little appreciated. Most important among these is an inherent dependence on species number. To calculate any of these indices it is necessary to know the number of species in the underlying sample universe or community; a number which, if not imaginary, is close to impossible to determine for most ecological applications. Workers wishing to calculate equitability have frequently selected an index and substituted the number of species in the sample ( $\hat{S}$ ) for the number in the universe ( $S$ ). There are two serious problems with this approach. Because  $S$  is always underestimated by  $\hat{S}$ , equitability is always overestimated. Secondly, differences in sample size or simple stochastic variation in  $\hat{S}$  can greatly affect the results obtained (29, 63, 71).

An example from Peet (manuscript) illustrates the sensitivity of these indices to the small sampling variation that can result when substituting the sample species number for the universe species number. Table 3 lists the responses of nine indices to a very minor change in sample composition. In a sample of 1000 individuals the species of one individual has been changed. Such a change can easily result from chance variation. If an index is to produce meaningful results it should be insensitive to this sort of variability. Examination of Table 3 reveals that the equitability indices tested are highly sensitive to this change. The only index for which this is not obvious is that of Buzas & Gibson (7). In this case it must be noted that the minimum possible value is 1.00, which implies a 26% change has occurred. The implication is clear. While it is legitimate to calculate the equitability of a sample, one cannot make inferences about the underlying community unless the total number of species in that community is known.

Fager (12) has recommended that all heterogeneity indices be scaled in the manner of an equitability index. Fager apparently assumed that scaling reduces the variation of heterogeneity indices with changing sample size. As Table 3 illustrates, the opposite is true; scaling increases variation with sample size because it compares a heterogeneity measure with a hypothetical maximum that is highly dependent on sample size. As with all simple equitability measures, this approach can only be applied when the number of species in the underlying universe is known.

### *Hill's Ratios*

An entirely different form of equitability measure has been proposed by Hill (19) based on his series of diversity numbers. Because all the diversity numbers in Hill's series have the same units, their results can be compared in the form of ratios. The different diversity numbers also have varying sensitivity to changes in rare and common species, as previously discussed. Consequently, variation in the values of the ratios expresses variation in the contribution of rare and common species to the measured heterogeneity. Consider three hypothetical communities represented by

Table 3 Equitability values for two similar samples

INDEX <sup>a</sup>	POPULATION A (500,300,200)	POPULATION B (500,299,200,1)
$V'H'$ (Pielou's $J'$ )	0.937	0.750
$V'H$ (Pielou's $J$ )	0.937	0.747
$VH'$	0.936	0.746
$V$ Simpson	0.930	0.827
Redundancy (Patten)	0.064	0.254
Standard Deviation (Fager)	0.735	0.583
Number of Moves (Fager)	0.699	0.466
$\text{Exp } (H')/S$ (Buzas & Gibson)	0.933	0.705
$H(S)/M(S)$ (Lloyd & Ghelardi) <sup>b</sup>	0.344	0.272

<sup>a</sup>  $V$  and  $V'$  notation follows Hurlbert (20).

<sup>b</sup> Lloyd & Ghelardi's alternative formulation of  $S'/S$  produces similar results.



dominance diversity curves E, F, and G in Figure 3. Curves E and G represent two alternative means of increasing the heterogeneity of community F. Community G has a lower concentration of dominance with constant species number, while community E has a larger number of rare species but the distribution of dominance among the most important species is kept constant. As shown in Table 2, the ratio of  $N_1$  to  $N_2$  ( $=R_{1,2}$ ) produces an increase for E relative to F, but a decrease for G compared to F. This ratio is thus an indicator of the shape of the underlying dominance diversity relationship and of the relative contribution of richness and equitability to the heterogeneity indices.

The Hill ratios depend only on the calculated diversity numbers. Thus for all diversity numbers except  $N_0$  the ratios are independent of sample size. While the previously discussed equitability measures can only be used when the number of species in the universe is known, Hill's ratios can be used any time heterogeneity can be calculated. These ratios do not, however, measure equitability as it is normally defined and it seems less ambiguous to refer to them simply as Hill's ratios.

Hill's ratios are not always easy to interpret. As heterogeneity decreases, the values of the diversity numbers and consequently the ratios, will converge toward one. Thus a low value of a ratio could either mean that the overall heterogeneity is low or that the dominance is spread over a number of the more common species in the community. Another difficulty is that dominance diversity curves with slopes approaching either zero or infinity will have ratios approaching one. For this reason curves H and I in Figure 4 both have ratios ( $R_{1,2}$ ) less than obtained for F, which has an intermediate slope. To date Hill's ratios have been little used but they offer a potential means of examining community structure once their response behavior is more fully understood.

## EXAMPLES

Two examples using field data illustrate the application of diversity indices. The first example compares the responses of four indices to two reasonably well-documented gradients in bird species diversity; increasing diversity with successional development and increasing diversity with decreasing latitude (Figure 5). The successional sequence is drawn from a paper by Kricher (26) utilizing averaged data for summer resident birds along a sere composed of two-year fields, old fields with cedar, and near-climax oak forest. The latitudinal gradient compares resident bird diversity for three hemlock-hardwood forests distributed along the Appalachians between New York and North Carolina. These data are taken from papers by Odum (47), Ken-deigh (24), and Stewart & Aldrich (66).

The index values (Figure 5) leave little doubt that bird species diversity increases across both gradients. Note the overall parallelism in behavior of the different indices to be expected when samples do not differ widely in dominance patterns. This observation is in agreement with that of Tramer (67) that the dominance patterns of temperate bird communities are reasonably uniform.

Also of interest is the response of the ratio  $R_{1,2}$ . For the successional sequence the ratio decreases steadily indicating a decrease in dominance through successional time of the one or two most important species. The latitudinal sequence does not

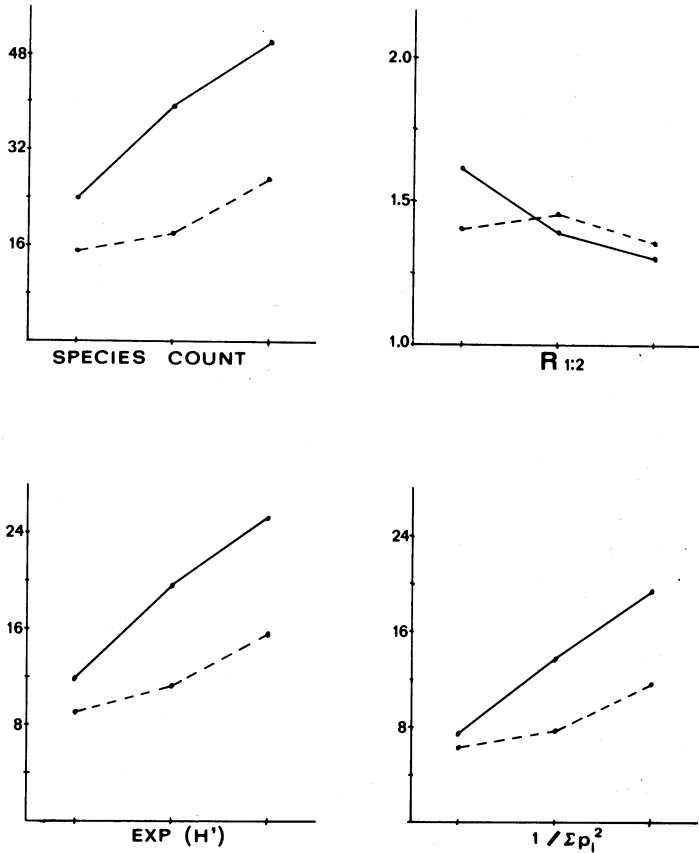


Figure 5 Values of four indices along two gradients in bird species diversity. The broken line represents breeding bird diversity at three points along a north to south latitudinal gradient of Appalachian hemlock-hardwood forests (data from Odum, 47). The solid line represents summer resident bird diversity at three points along a successional sequence in New Jersey (data from Kricher, 26).

show an obvious trend in the ratio, suggesting that the species-abundance relationship does not change greatly with latitude for mature communities.

A more complex case is that of Peet (unpublished data) describing the plant communities of Rocky Mountain National Park, Colorado. Five composite altitudinal samples are considered (Figure 6). These represent sequentially: *Picea engelmannii*-*Abies lasiocarpa* krummholz, *Picea*-*Abies* forest, *Pinus contorta* forest, *Pinus ponderosa*-*Pseudotsuga* forest, and *Pinus ponderosa* woodland. In all cases the indices present a U-shaped response reflecting the floristic poverty of the dense, middle elevation, coniferous forests. Interestingly, the heterogeneity indices show a

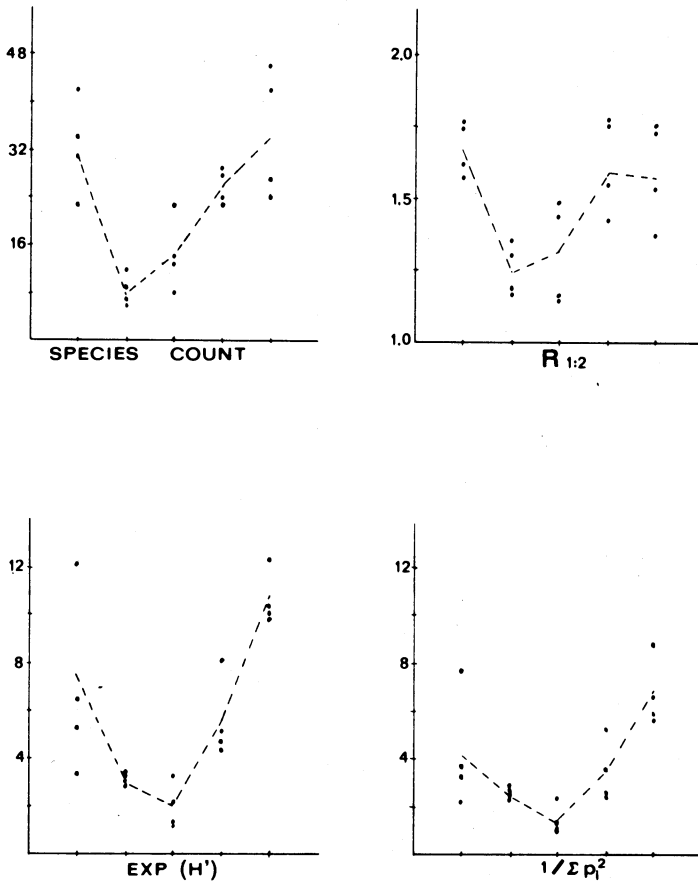


Figure 6 Values of four diversity indices along a complex altitudinal-physiognomic gradient of plant communities in Rocky Mountain National Park, Colorado. The five composite altitudinal levels represent respectively: open *Picea engelmannii*-*Abies lasiocarpa* krummholz, *Picea*-*Abies* forest, *Pinus contorta* forest, *Pinus ponderosa*-*Pseudotsuga* forest, and open *Pinus ponderosa* woodland.

minimum in the *Pinus contorta* stands, but the *Picea*-*Abies* stands have a lower species richness, as indicated by species per 0.1 hectare sample. The type II index shows a greater dip for the *Pinus contorta* stand than the type I because of its greater sensitivity to the high *Pinus* dominance. In contrast, the type I index shows the greater response at the ends of the gradient due to the increased numbers of rare species. The first of these differences is also reflected in the Hill's ratio  $R_{1/2}$ , which, despite depression due to low heterogeneity, averages higher for the *Pinus contorta* forest than for the *Picea*-*Abies* forest.

## CONCLUSION AND SUMMARY

A number of concepts are lumped under the title of diversity. Of particular importance are species richness or the number of species in the community, and equitability or the evenness with which importance is distributed among the species. When the form of the underlying species-abundance relation is known, it is possible to estimate both richness and equitability based on this known functional relationship. When both the underlying species-abundance relation and the number of species in the sampling universe are unknown, neither richness nor equitability can be estimated. Rather, a third concept, called heterogeneity, which combines the richness and equitability measures, must be used.

Heterogeneity indices can be divided into two types. Type I indices are most sensitive to changes in the rare species and are characterized by response functions, the second derivatives of which increase with decreasing species importance. Type II indices are most sensitive to changes in the most abundant species and have second derivatives that are constant or decrease as importance approaches zero. The best known examples are the Shannon formula, which is of type I, and the Gini index (complemented Simpson index), which is of type II. Conceivably a third class most sensitive to species of intermediate importance might be added.

The characteristics of diversity indices and their responses to changes in community structure should be understood as the basis for their use. An investigator using the approach presented in this paper should be able to select or design an index emphasizing that aspect of diversity he is most interested in measuring. Conclusions regarding diversity measurements that seem of major usefulness are summarized in the following outline.

- A. Richness indices, based directly on species number
  1. Species number per sample measures richness as here defined and is the most basic and general diversity measurement. It is, however, affected by arbitrary choice of sample size and potential error in determining the number of species.
  2. The rate of increase in number of species with increase in sample size can be measured in one of several forms. Such a measure escapes the arbitrary choice of a standard sample size, but depends on knowledge of the underlying relation of species number to sample size and the assumption that their relation remains constant among the samples being compared.
- B. Heterogeneity indices, based on a combination of richness and equitability
  1. An infinite number of potential indices is available. The response curve formed by graphing the component of such an index which is summed over all species can provide basic information on the effects of different types of changes in sample composition. Response curves can also be used to distinguish different types of heterogeneity indices.
  2. Type I heterogeneity indices are defined as being most sensitive to changes in the importance of the rare species in the sample. The most frequently encountered example is the Shannon formula. Use of the exponentiated form of the Shannon index is suggested for interpretational reasons.

3. Type II heterogeneity indices are most sensitive to changes in the most common species. The best known example is the Gini or complemented Simpson index. The reciprocal of Simpson's index is suggested for general application.
- C. Equitability indices, based on the evenness of the distribution of importance between species
  1. Use of simple scaled heterogeneity indices or related indices (such as Pielou's  $J$  or Lloyd & Ghelardi's  $\epsilon$ ) to measure equitability is not possible unless the total number of species in the sampling universe is known, a rare occurrence in ecology. In general, the use of scaled indices is not recommended.
  2. Hill's ratios provide a new and potentially useful method of examining community diversity, but this approach has not yet been fully explored.

#### ACKNOWLEDGMENTS

The author gratefully acknowledges the advice and encouragement of Robert H. Whittaker and Orie L. Loucks. Many helpful comments on the manuscript were offered by David C. Lewin, Robert W. Poole, and Simon Levin.

#### Literature Cited

1. Auclair, A. N., Goff, F. G. 1971. Diversity relations of upland forests in the western Great Lakes area. *Am. Natur.* 105:499-528
2. Austin, M. P. 1968. Relationships among functional properties of Californian grassland. *Nature* 217:1163
3. Berger, W. H., Parker, F. L. 1970. Diversity of planktonic Foraminifera in deep-sea sediments. *Science* 168: 1345-47
4. Bhargava, T. N., Doyle, P. H. 1974. A geometric study of diversity. *J. Theor. Biol.* 43:241-51
5. Bowman, K. Q., et al 1969. Comments on the distribution of indices of diversity. *Proc. Int. Symp. Stat. Ecol.* 3:315-59
6. Bullock, J. A. 1971. The investigation of samples containing many species. I. Sample description. *Biol. J. Linn. Soc.* 3:1-21
7. Buzas, M. A., Gibson, T. G. 1969. Species diversity: benthonic Foraminifera in the western North Atlantic. *Science* 163:72-75
8. Crow, J. F., Kimura, M. 1970. *An Introduction to Population Genetics Theory*. New York: Harper & Row. 591 pp.
9. Dickman, M. 1968. Some indices of diversity. *Ecology* 49:1191-93
10. Eberhardt, L. L. 1969. Some aspects of species diversity models. *Ecology* 50:503-5
11. Edden, A. C. 1971. A measure of species diversity related to the lognormal distribution of individuals among species. *J. Exp. Mar. Biol. Ecol.* 6:199-209
12. Fager, E. W. 1972. Diversity: a sampling study. *Am. Natur.* 106:293-310
13. Fisher, R. A., Corbet, A. S., Williams, C. B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12:42-58.
14. Gini, C. 1912. Variabilità e mutabilità. *Studi Economico-Giuridici Fac. Giurisprudenza Univ. Cagliari*, A. III, parte II.
15. Good, I. J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237-64
16. Greenberg, J. H. 1956. The measurement of linguistic diversity. *Language* 32:109-15
17. Hairston, N. G., Byers, G. W. 1954. The soil arthropods of a field in southern Michigan. A study in community ecology. *Contrib. Lab. Vertebrate Biol., Univ. Mich.* 64:1-37
18. Hendrickson, J. A. Jr., Ehrlich, P. R. 1971. An expanded concept of "species diversity." *Notulae Natur. Acad. Natur. Sci. Philadelphia* 439:1-6
19. Hill, M. O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427-32
20. Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577-86

21. Hutcheson, K. 1970. A test for comparing diversities based on the Shannon Formula. *J. Theor. Biol.* 29:151-54
22. Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? *Am. Natur.* 93:145-59
23. Johnson, M. P., Raven, P. H. 1970. Natural regulation of plant species diversity. *Evol. Biol.* 4:127-62
24. Kendeigh, S. C. 1946. Breeding birds of the beech-maple-hemlock community. *Ecology* 27:226-45
25. Kohn, A. J. 1968. Microhabitats, abundance and food of *Conus* on atoll reefs in the Maldive and Chagos Islands. *Ecology* 49:1046-61
26. Kricher, J. C. 1973. Summer bird species diversity in relation to secondary succession on the New Jersey piedmont. *Am. Midl. Natur.* 89:121-37
27. Leti, G. 1965. Sull' entropia, su un indice del Gini e su altre misure dell' eterogeneità di un collettivo. *Metron* 24:332-78 [English Summary]
28. Lieberman, S. 1964. An extension of Greenberg's linguistic diversity measures. *Language* 40:526-31
29. Lieberman, S. 1969. Measuring population diversity. *Am. Soc. Rev.* 34:850-62
30. Lloyd, M. 1964. Weighting individuals by reproductive value in calculating species diversity. *Am. Natur.* 98:190-92
31. Lloyd, M., Ghelardi, R. J. 1964. A table for calculating the 'equitability' component of species diversity. *J. Anim. Ecol.* 33:217-25
32. Lloyd, M., Zar, J. H., Karr, J. R. 1968. On the calculation of information-theoretical measures of diversity. *Am. Midl. Natur.* 79:257-72
33. Longuet-Higgins, M. S. 1971. On the Shannon-Weaver index of diversity, in relation to the distribution of species in bird censuses. *Theor. Pop. Biol.* 2:271-89
34. MacArthur, R. H. 1957. On the relative abundance of bird species. *Proc. Nat. Acad. Sci. USA* 43:293-95
35. MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. *Ecology* 39:599-619
36. MacArthur, R. H. 1965. Patterns of species diversity. *Biol. Rev.* 40:510-33
37. MacArthur, R. H. 1972. *Geographical Ecology*. New York: Harper & Row. 269 pp.
38. MacArthur, R. H., MacArthur, J. W. 1961. On bird species diversity. *Ecology* 42:594-98
39. MacArthur, R. H., Wilson, E. O. 1967. *The Theory of Island Biogeography*. Princeton, NJ: Princeton Univ. Press. 203 pp.
40. McIntosh, R. P. 1967. An index of diversity and the relation of certain concepts to diversity. *Ecology* 48:392-404
41. Margalef, R. 1951. Diversidad de especies en las comunidades naturales. *Publ. Inst. Biol. Apl. Barcelona* 9:5-27 [English Summary]
42. Margalef, R. 1957. La teoria de la informacion en ecologia. *Mem. Real Acad. Cienc. Artes Barcelona* 32:373-449 [Transl. in *Gen. Syst.* 3:36-71]
43. Margalef, R. 1969. Diversity and stability: a practical proposal and a model of interdependence. *Brookhaven Symp. Biol.* 22:25-37
44. Mason, H. L., Langenheim, J. H. 1957. Language analysis and the concept *environment*. *Ecology* 38:325-40
45. Menhinick, E. F. 1964. A comparison of some species-individuals diversity indices applied to samples of field insects. *Ecology* 45:859-61
46. Monk, C. D. 1967. Tree species diversity in the eastern deciduous forest with particular reference to north central Florida. *Am. Natur.* 101:173-87
47. Odum, E. P. 1950. Bird populations of the Highlands (North Carolina) Plateau in relation to plant succession and avian invasion. *Ecology* 31:587-605
48. Odum, H. T., Cantlon, J. E., Kornicker, L. S. 1960. An organization hierarchy postulate for the interpretation of species-individual distributions, species entropy, ecosystem evolution, and the meaning of a species-variety index. *Ecology* 41:395-99
49. Patten, B. C. 1962. Species diversity in net phytoplankton of Raritan Bay. *J. Mar. Res.* 20:57-75
50. Pielou, E. C. 1966. Species-diversity, and pattern-diversity in the study of ecological succession. *J. Theor. Biol.* 10:370-83
51. Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* 13:131-44
52. Pielou, E. C. 1966. Shannon's formula as a measure of species diversity: its use and misuse. *Am. Natur.* 100:463-65
53. Pielou, E. C. 1967. The use of information theory in the study of the diversity of biological populations. *Proc. Berkeley Symp. Math. Stat. Prob.*, 5th 4:163-77
54. Pielou, E. C. 1969. *An Introduction to Mathematical Ecology*. New York: Wiley-Interscience. 286 pp.

55. Poole, R. W. 1974. *An Introduction to Quantitative Ecology*. New York: McGraw. 532 pp.
56. Preston, F. W. 1948. The commonness, and rarity, of species. *Ecology* 29:254-83
57. Preston, F. W. 1960. Time and space and the variation of species. *Ecology* 41:611-27
58. Preston, F. W. 1962. The canonical distribution of commonness and rarity. *Ecology* 43:185-215, 410-32
59. Preston, F. W. 1969. Diversity and stability in the biological world. *Brookhaven Symp. Biol.* 22:1-12
60. Sager, P. E., Hasler, A. D. 1969. Species diversity in lacustrine phytoplankton. I. The components of the index of diversity from Shannon's formula. *Am. Natur.* 103:51-59
61. Sanders, H. L. 1968. Marine benthic diversity: a comparative study. *Am. Natur.* 102:243-82
62. Shannon, C. E., Weaver, W. 1949. *The Mathematical Theory of Communication*. Urbana, Ill.: Univ. Illinois Press. 117 pp.
63. Sheldon, A. L. 1969. Equitability indices: dependence on the species count. *Ecology* 50:466-67
64. Simberloff, D. 1972. Properties of the rarefaction diversity measurement. *Am. Natur.* 106:414-18
65. Simpson, E. H. 1949. Measurement of diversity. *Nature* 163:688
66. Stewart, R. E., Aldrich, J. W. 1949. Breeding bird populations in the spruce region of the central Appalachians. *Ecology* 30:75-82
67. Tramer, E. J. 1969. Bird species diversity: components of Shannon's formula. *Ecology* 50:927-29
68. Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* 30:279-338
69. Whittaker, R. H. 1965. Dominance and diversity in land plant communities. *Science* 147:250-60
70. Whittaker, R. H. 1969. Evolution of diversity in plant communities. *Brookhaven Symp. Biol.* 22:178-96
71. Whittaker, R. H. 1972. Evolution and measurement of species diversity. *Taxon* 21:213-51
72. Whittaker, R. H., Woodwell, G. M. 1969. Structure, production and diversity of the oak-pine forest at Brookhaven, New York. *J. Ecol.* 57:155-74
73. Williams, C. B. 1964. *Patterns in the Balance of Nature*. London: Academic. 324 pp.
74. Williamson, M. 1973. Species diversity in ecological communities. In *The Mathematical Theory of the Dynamics of Biological Populations*, ed. M. S. Bartlett, R. W. Hiorns, 325-35. London: Academic
75. Woodwell, G. M. 1967. Radiation and the patterns of nature. *Science* 156:461-70
76. Yule, G. U. 1944. *The Statistical Study of Literary Vocabulary*. London: Cambridge Univ. Press. 306 pp.