A CASE STUDY IN CONCEPT DETERMINATION: ECOLOGICAL DIVERSITY

James Justus

1 INTRODUCTION

Some biological communities are more complicated than others. For example, tropical communities usually contain more species [Pianka, 1966; Willig *et al.*, 2003], there is evidence their species interact more intensely [Janzen, 1970; Møller, 1998], these interactions are more variegated in form [Dyer and Coley, 2001], and they exhibit more trophic levels than high latitude communities [Oksanen *et al.*, 1981; Fretwell, 1987]. Ecologists often use the concept of diversity to represent differences in the "complicatedness" of communities: tropical communities are often said to be more ecologically diverse than tundra communities.

At a coarse level of description, the vague connotation accompanying the term 'diversity' adequately captures the imprecise judgments that some communities are more complicated than others. Disagreement arises, however, over how the concept should be operationalized. As early as 1969, Eberhardt [1969, p. 503] characterized the ecological literature on diversity as a "considerable confusion of concepts, definitions, models, and measures (or indices)." A few years later, Hurlbert [1971, p. 577] argued that, "the term 'species diversity' has been defined in such various and disparate ways that it now conveys no information other than 'something to do with community structure'." MacArthur [1972, p. 197] similarly suggested that the term 'diversity' should be excised from ecological vocabulary as doing more harm than good, and that ecologists had, "wasted a great deal of time in polemics about whether [Simpson's] or [Shannon's] or $\frac{N!}{N_1!N_2!...N_n!}$ or some other measure [of diversity] is 'best'."[1] As these remarks indicate, ecologists have proposed several mathematical measures that differ about what properties are given priority over others in assessing diversity and which differ in mathematical form. Disagreements about these issues raise the question of what properties of a community *should* be considered part of its diversity and, in turn, what adequacy conditions the concept should satisfy.

---

[1] See sections 3 and 4 for a discussion of these diversity measures.

Handbook of the Philosophy of Science. Volume 11: Philosophy of Ecology.
Volume editors: Kevin deLaplante, Bryson Brown, and Kent A. Peacock. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Section 2 describes and defends seven adequacy criteria for the concept of ecological diversity. It also argues two additional criteria found in the ecological literature are untenable. The primary focus is adequacy criteria for measures of diversity, such as Shannon's and Simpson's, that make no assumption about the underlying distribution of individual organisms among species in a community. For this reason, these indices are sometimes called *nonparametric* (e.g. [Lande, 1996]) to distinguish them from indices derived from parameters of statistical models of species abundance, such as the log series [Fisher *et al.*, 1943] and log normal [Preston, 1948], or from biological models, such as the broken stick and overlapping niche model [MacArthur, 1957].[2] Unlike parametric indices, nonparametric diversity indices are applicable to any biological community with any species abundance distribution.[3] This analysis also assumes that communities have been exhaustively sampled, thereby avoiding complex issues about the adequacy of diversity indices given imperfect and incomplete sampling to focus on the problem of specifying the concept of ecological diversity when complete knowledge about the community's relevant properties is available.[4]

Like most ecological literature on the concept of diversity, the focus is on species richness and evenness as components of diversity, although issues about how other information (e.g. taxonomic information) should affect assessments of diversity are occasionally touched upon. A myriad of indices combine species richness and evenness into a single measure of diversity, the two most popular being Simpson's and Shannon's. Sections 3 and 4 describe these indices and evaluate how they fare against the adequacy criteria defended in Section 2. Despite its greater popularity, Shannon's index performs worse than Simpson's. Section 5 concludes by assessing an influential criticism of the role of the diversity concept within ecology.

2 ADEQUACY CRITERIA FOR THE CONCEPT OF ECOLOGICAL DIVERSITY

Like most systems studied in science, biological communities can be represented with different degrees of specificity. With low specificity, a community can be represented simply in terms of the species it contains and how individual organisms of the community are distributed among these species.[5] This information is provided

---

[2] See [Preston, 1962a; 1962b; May, 1975; Rosenzweig, 1995] for reviews.

[3] In fact, Lande [1996, p. 5] suggests that being nonparametric, and thus applicable to all biological communities, is a defensible adequacy condition for a diversity index. Some ecologists have also criticized that there is no theoretical justification for statistical models of species abundance distribution, and only poor ones for most biological models [Krebs, 1989, Ch. 10].

[4] See [Horn, 1966; Pielou, 1975; 1977; Patil and Tallie, 1982a] for extensive discussions of these issues.

[5] The target of this analysis is therefore quite modest. Greater representational specificity is achieved, for example, if interactions between species in a community are described with differential or difference equations in addition to how individuals are distributed among species (see [Justus, 2006]). For a much broader representational scope on idversity, see Maclaurin and Sterelny's [2008] analysis of biodiversity.

by the *proportional species abundance vector*, V_p, of a community:

(1) $V_p = \langle p_1, ..., p_i, ..., p_n \rangle$;

in which n designates the number of species in the community, i.e. its *species richness*; p_i designates the proportional abundance of the i-th species in the community; the p_i are ordered from most to least abundant (ties broken by random selection), i.e. $p_1 \geq ... \geq p_i \geq ... \geq p_n$; and $\sum\limits_{i=1}^{n} p_i = 1$. The only properties of species V_p represents are their proportional abundances. Functional, trophic, and taxonomic differences (besides the species level) are not represented. Proportional abundances of species in a community often change over time for a variety of reasons (*e.g.* migration, interspecific interactions such as predation, competition, *etc.*) so V_p must be updated as communities change.

'Abundance' is an ambiguous term. Besides referring to the number of individual organisms of a species (a discrete quantity), it can also refer to their biomass (a continuous quantity). Accordingly, p_i can designate either: (i) the proportion of individuals of species i in a community given by $\dfrac{N_i}{N}$ where N_i is the number of individuals of species i and N is the total number of community individuals, such as the proportion of wolves in an community; or, (ii) species i's proportion of total community biomass, such as dry weight of a particular plant species in a forest community. p_i may differ significantly on these two interpretations, so ideally V_p should be calculated according to both interpretations for a given biological community. If it is unclear how to count individual organisms, as is the case for some clonal plant species or asexually reproducing marine species, the biomass interpretation of p_i is preferable.

Mathematically, components of V_p can take values of zero to represent species with zero abundance. Unlike species for which $p_i > 0$, adding or subtracting these terms from V_p need not change the other proportional abundances to ensure $\sum\limits_{i=1}^{n} p_i = 1$. Biologically, however, species for which $p_i = 0$ cannot be part of the community represented by V_p. To be one of the species comprising a community, the community must contain at least one representative of that species. As a biological collection, to deny this stipulation for communities would require commitment to the idea that a community can be represented to contain a species not instantiated by any of its members. Depending on the interpretation of p_i (see above), $p_i \geq \frac{1}{N}$ or $p_i \geq \frac{b_i}{B}$ for all i is assumed where N designates the total number of individuals in the community, B designates the total community biomass, and b_i designates the minimum biomass of an individual of species i.[6] In modeling contexts with different goals, such as in studies of extinction and migration processes, it may be useful to allow zero V_p components to represent when species have gone locally extinct or have emigrated completely from a community. Once a species

---

[6] For expositional convenience, only interpretation (i) from above will be discussed in the following unless specified otherwise.

disappears from a community, however, it is no longer part of that community and does not contribute to its diversity.[7]

Ecologists widely agree that two properties of a community should be part of its diversity: *species richness* and *evenness* [Pielou, 1966; 1975; 1977; Tramer, 1969; Patil and Taillie, 1982a; 1982b; Margurann, 1988; 2004].[8] Consider two simple communities, A and B, both composed of two species s_1 and s_2. A and B have the same species richness (two). If the proportions of individuals distributed among the two species are $p_1 = 0.02\%$ and $p_2 = 99.98\%$ for A and $p_1 = 50\%$ and $p_2 = 50\%$ for B, B is said to be more even than A.

The widespread beliefs that species richness and evenness are components of diversity reflect intuitive constraints on the concept that can be formulated as explicit adequacy conditions:

(A1) for a given evenness, diversity should increase as species richness increases (i.e. as n from [1] above increases); and,

(A2) for a given species richness, diversity should increase as evenness increases.

Note that neither (A1) nor (A2) necessitate a particular mathematical form to the increase in diversity required.

The first condition codifies an incontestable feature of the diversity concept: the diversity of a collection increases as the number of different types of entities in the collection increases. Applied to a biological collection such as a community, (A1) therefore captures the intuitive idea that a community composed of one thousand species is more diverse than one composed of ten.

But there is a difficulty with (A1) as formulated. Unlike the clause “for a given species richness” in (A2), (A1) contains a qualification, “for a given evenness,” for which V_p does not provide a quantitative characterization.[9] The problem is that increases in species richness (represented by new p_i) necessitate changes in the p_i comprising V_p since all the p_i must sum to one following any change in species richness. These changes do not *necessitate* a change in evenness, but the absence of a quantitative characterization makes it unclear how evenness can remain static as species richness changes.[10] To avoid this difficulty, (A1) is often reformulated as:

---

[7] Alternatively, the stipulation against zero p_i could be replaced with the proviso that species richness is determined by the number of *nonzero* p_i in V_p and that only they determine community diversity.

[8] McIntosh [1967] was probably the first to coin the term ‘species richness’ to refer to the number of species in a community. ‘Evenness’ and ‘equitability’ are used interchangeably in the ecological literature (e.g. [Lloyd and Ghelardi, 1964; McIntosh, 1967; Tramer, 1969; Peet, 1974; 1975]).

[9] n from (1) provides a quantitative characterization of species richness.

[10] Similarly, although the simple examples discussed above and below provide an informal grasp of how communities can differ in evenness, absence of a quantitative characterization also makes the clause “as evenness increases” of (A2) unclear. The remainder of this section proposes adequacy conditions which help precisify the evenness concept.

(A1′) of two maximally even communities, the more species rich community is more diverse [Pielou, 1975, p. 7].

A quantitative characterization of 'maximally even' is provided by a third adequacy condition (A3) below.

Beyond its intuitive appeal, it is worth pausing over the reason (A2) should be accepted. Consider two communities, each composed of 100 species and 10000 total individual organisms. A community in which there are 100 individuals of each species seems more diverse than one with 9901 individuals of one species and one individual each of the other 99. The reason seems to be that besides a consideration of the number of types of entities in a collection, diversity also involves a consideration of how well they are represented. For this reason, diversity is often interpreted as the *apparent* or *effective* number of species present in a community (e.g. [Hill, 1973; Peet, 1974]). For example, to an observer with imperfect faculties of perception, or an ecologist with insufficient field time, or employing sampling methods with unavoidable limitations, the first community with evenly distributed individuals will usually appear to contain more species than the second community, despite their identical richness.[11] In this way, (A2) captures the intuitive idea that community B is more diverse than community A from above.

Thus far, evenness has not been explicitly characterized and (A1) and (A2) place no constraints on the concept. For a given species richness, evenness is clearly maximized when individuals of the community are equally distributed among species, i.e. when $p_i = \frac{1}{n}$ for all i. This constraint corresponds to another adequacy condition:

(A3) for a given species richness, diversity is maximal when individuals of the community are distributed equally among species (*i.e.* when evenness is maximal).

Let $V_p^{\max}$ designate the maximally even proportional species abundance vector for a given richness.

Similarly, evenness is clearly minimized when community individuals are maximally unequally distributed. Specifically, diversity is minimal when all but $(n-1)$ of the individual organisms comprising the community are of one species and the rest are equally distributed (one each) among the other $(n-1)$ species, i.e. when $p_1 = \frac{N-(n-1)}{N}$.[12] Formulated as an explicit adequacy condition:

(A4) for a given species richness, diversity is minimal when individuals of the community are distributed maximally unequally among species (i.e. when evenness is minimal).

---

[11] A different line of thought also motivates (A2). A biological community is a set of organisms of different species. Sets are characterized by properties of their members. Members of an uneven community poorly represent some species, while each species of an even community is equally represented by its members. As a set, the characterization of a community with evenly distributed individuals therefore depends more significantly on a greater number of species-types than an uneven community.

[12] Recall that n designates the number of species and N designates the total number of individual organisms.

Let V_p^{min} designate the minimally even proportional species abundance vector for a given richness.

Adequacy conditions (A1)–(A4) are found throughout the ecological literature [Hill, 1973; Pielou, 1975; 1977; Magurran, 1988; 2004; Lande, 1996; Sarkar, 2007]. Building on (A2)–(A4), a further constraint on the evenness concept, and thus on diversity, can be formulated. Focusing on (A3),[13] if evenness is maximal for V_p^{max}, evenness must decrease as V_p diverges from it. This decrease can be quantified in many ways, but one rationale for doing so restricts the range of possible methods of quantification. Recall that the only differences between species being considered are their proportional abundances; V_p does not represent taxonomic, trophic, functional, and other interspecific differences. Besides their proportional abundances, different species are therefore treated as equally important in assessing the diversity of a community. Thus, if evenness decreases because one species deviates from its maximally even proportional abundance $(\frac{1}{n})$, an equal deviation from the maximally even proportional abundance by another species should induce an identical decrease in evenness and thus in diversity. Formulated as an explicit adequacy condition:

(A5) for a given species richness, if V_p^i and V_p^j are proportional species abundance vectors that deviate from V_p^{max} because species i and species j, respectively, deviate equally from a $\frac{1}{n}$ proportional abundance, evenness decreases by the same amount in both cases.

Put informally, (A5) stipulates that assessment of diversity is blind to species identity. It thereby captures the frequently made assumption that evaluating community diversity requires treating species as equals in the absence of taxonomic, functional, or other data [Magurran, 2004, p. 11]. In such cases, only the extent a species' proportional abundance deviates, not what species it is, is relevant when assessing a community's diversity.[14] (A5) is neutral, however, about whether rare or abundant species are more important to the diversity of a community. It requires merely that *equal* changes in the abundances of two species from maximal evenness necessitate *equal* decreases in diversity.

Even with this further constraint, (A1)–(A5) are weak adequacy conditions in the sense that they do not determine a unique quantitative measure of diversity. In fact, most common quantitative indices satisfy them (see §§3–5). Distinct quantitative diversity indices result from different ways of integrating and quantifying species richness and evenness consistent with (A1)–(A5). Before discussing the

---

[13] Similar reasoning applies for (A4).

[14] Strictly speaking, (A5) follows from the way V_p was constructed. Recall that the p_i are ordered from most to least abundant. This was intended to impose a nonarbitrary ordering on the components of V_p. It also entails, however, that V_p^i and V_p^j referred to in (A5) are identical because p_i and p_j would fall at the same place in the ordering for V_p^i and V_p^j, respectively. If the ordering constraint were not imposed and species were assigned indices in V_p prior to determination of their proportional abundances, (A5) would constitute an independent requirement on diversity. As is, (A5) is retained to make the requirement explicit. I owe Samir Okasha for this clarification.

two most common such indices in the next sections, it is therefore important to consider whether any other defensible adequacy conditions would necessitate a particular quantitative index of diversity.

Notice that (A5) does not entail different *types* of deviations from $V_p^{\max}$ (or $V_p^{\min}$), such as those involving different numbers of species, must be accorded the same import for diversity. Of course, unequal deviations *of the same type* should necessitate different values of diversity. Consider, for example, the type of deviation in which one species i deviates from $\frac{1}{n}$. If p_i decreases from $\frac{1}{n}$ to $\frac{1}{2n}$, diversity should decrease less than if p_i decreases to $\frac{1}{3n}$ all else being equal because the decrease in evenness is greater in the latter case (see [A2]). If p_i decreases from $\frac{1}{n}$ to $\frac{1}{4n}$, however, (A5) entails nothing about whether diversity should decrease more or less than in a case in which p_i decreases from $\frac{1}{n}$ to $\frac{1}{2n}$ and some other p_k decreases from $\frac{1}{n}$ to $\frac{1}{2n}$.

What is needed is an method for evaluating evenness that would adjudicate between different types of deviations from $V_p^{\max}$ (or $V_p^{\min}$) for a given species richness. One natural method for doing so simply evaluates evenness in terms of the *distance* between V_p and $V_p^{\max}$. In general, a function $d : G \times G \rightarrow \mathbf{R}$ is a distance metric if it possesses three properties for all $x, y, z \in G$:

(P1) $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$;

(P2) $d(x, y) = d(y, x)$ (symmetry); and,

(P3) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality). [Kaplansky, 1977].

An infinite number of different functions satisfy these conditions and could therefore be used to measure deviation of V_p from $V_p^{\max}$. For instance, an especially simplistic distance metric satisfying (P1)–(P3) is $d(x,y) = 0$ if $x = y$, and 1 otherwise. This is plainly inappropriate as a metric for measuring the distance between biological communities represented by V_p and $V_p^{\max}$. According to this metric, communities A and B from above are at the same distance from $V_p^{\max}$.

Since components of V_p and $V_p^{\max}$ take real values, an appropriate distance metric, and certainly the most common, is the Euclidean metric:

$$(2) \quad d(V_p, V_p^{\max}) = \sqrt{\sum_{i=1}^{n} \left(p_i - \frac{1}{n}\right)^2}.^{15}$$

Measured in this way, the idea is that evenness is inversely related to the Euclidean distance between its species abundance vector and $V_p^{\max}$.[16] For example, if the diversity of $V_p^{\max}$ and $V_p^{\min}$ for a given richness are set at 1 and 0, respectively (see [A3] and [A4] above), the diversity of a community represented by V_p would take values on [0,1] determined by the Euclidean distance between V_p and $V_p^{\max}$.

---

[15] Note that I am not claiming the Euclidean metric is *uniquely* defensible.

[16] Similarly, since evenness is inversely related to the Euclidean distance, evenness is directly related to the Euclidean distance from $V_p^{\min}$.

Species abundance vectors that deviate from $V_p^{\max}$ in different ways but at the same distance from it would thereby have the same evenness; those at different distances from $V_p^{\max}$ would differ in evenness. In particular, for the two species abundance vectors discussed above, one in which p_i for one species decreases from $\frac{1}{n}$ to $\frac{1}{4n}$ and another in which p_i and p_k for two species decrease from $\frac{1}{n}$ to $\frac{1}{2n}$, the latter would be accorded greater evenness because its distance from $V_p^{\max}$ is smaller than the former. Codified in an explicit adequacy condition:

(A6) for a given species richness: (i) evenness decreases (increases) as the Euclidean distance from $V_p^{\max}$ ($V_p^{\min}$) increases; and, (ii) communities represented by species abundance vectors at the same Euclidean distance from $V_p^{\max}$ (or $V_p^{\min}$) have the same evenness.[17]

Similar to the way (A5) stipulates that diversity is blind to species identity, (A6) stipulates that diversity is blind to the type of deviation from $V_p^{\max}$ (or $V_p^{\min}$). (A6) does not necessitate a unique quantification of evenness because the decrease (or increase) required in clause (i) may take many mathematical forms (*e.g.* concave vs. convex, linear vs. nonlinear, exponential vs. nonexponential, *etc.*). Note that (A6) holds only if (A5) does as well.

Together with (A2), (A6) imposes a significant constraint on the diversity concept. It requires treating changes in the proportional abundances of rare and common species as equally important to diversity. According to (A6), for instance, diversity must decrease the same amount with a decrease in p_i for an extremely rare species and with an identical decrease in a much more common species.[18] Thus, (A6) precludes diversity from being partially sensitive to the proportional abundances of rare or common species. In particular, it requires that species abundance vectors in which several species are very abundant and a few are very rare, and in which several species are very rare and a few are very abundant have the same diversity if their distances from $V_p^{\max}$ are identical. (A6) thereby captures the same idea underlying (A5): that diversity requires treating all species as equals. To illustrate, consider two species abundance vectors $\left\langle \frac{8}{15}, \frac{4}{15}, \frac{3}{15} \right\rangle$ and $\left\langle \frac{7}{15}, \frac{6}{15}, \frac{2}{15} \right\rangle$. The first contains one abundant species and two rarer species (the maximally even p_i for each species is $\frac{1}{3}$). The second contains two abundant species and one rare species. The distances between each vector and $V_p^{\max}$ are identical and hence both are accorded the same diversity by (A6).

For the same reason diversity should be blind to species identity (see [A5]) and blind to the type of deviation from $V_p^{\max}$ (or $V_p^{\min}$) (see [A6]), diversity should

---

[17] Since Euclidean distance is not the only defensible distance metric (see footnote above), (A6) need not be formulated with it. I focus on Euclidean distance because it is the most common metric for vectors with real components and is clearly defensible. Smith and Wilson [1996], for instance, describe several desirable properties of potential indices of evenness captured by (A6). However, the degree of agreement between analyses of diversity indices based on formulations of (A6) and (A7) with different defensible distance metrics is currently unknown.

[18] This holds when all species have proportional abundances $< \frac{1}{n}$. For species with proportional abundances $> \frac{1}{n}$, evenness will increase as their proportional abundances decrease towards $\frac{1}{n}$.

not be partial to particular distances between V_p and $V_p^{\max}$ (or $V_p^{\min}$) in the sense that diversity should decrease *uniformly* (i.e. linearly) as the distance between V_p and $V_p^{\max}$ increases. Formulated as an explicit adequacy condition:

(A7) for a given species richness, evenness decreases (increases) *linearly* as the Euclidean distance between V_p and $V_p^{\max}$ (or $V_p^{\min}$) increases.

(A7) requires equal intervals of distance (as measured by equation [2]) correspond to equal differences in diversity values, regardless of the specific distance V_p is from $V_p^{\max}$. Specifically, if $d(V_p^i, V_p^{\max}) = x$ and the difference in diversity value between V_p^i and $V_p^{\max}$ is y, then if $d(V_p^j, V_p^{\max}) = x$ also, the difference in diversity value between V_p^j and $V_p^{\max}$ is also y. Note that (A7) holds only if (A6) does as well.

(A6) and (A7) both follow from a principle sometimes mentioned in discussions of ecological diversity (e.g. [Krebs, 1989; Magurran, 2004]). The principle is that diversity should not be partial among *individual organisms*, just as it should not be partial among *species* in a community. Specifically, in assessments of diversity in the absence of taxonomic, functional, and other types of information, individual organisms should contribute to diversity in proportion only to the proportional abundance of the species to which they belong. If different types of deviations from $V_p^{\max}$ are weighted differently than as dictated by equation (2), i.e. (A6) is violated, some individuals will contribute more (or less) to diversity merely because they are a member of a species that has deviated from $\frac{1}{n}$ in a way favored (or disfavored) by the candidate diversity index. Similarly, if different distances from $V_p^{\max}$ are weighted differently in assessing diversity, i.e. (A7) is violated, some individuals will contribute more (or less) to diversity merely because they are a member of a species with a proportional abundance at a distance from $\frac{1}{n}$ that is favored (or disfavored) by the candidate index. In either case, individual organisms would not be treated as equals in determining the diversity of a community composed of them.[19]

Before evaluating Shannon and Simpson's indices against (A1)–(A7) in the next section, this section concludes by considering another adequacy condition for diversity proposed by Lewontin [1972] and recently endorsed by Lande [1996].[20] It concerns the relationship between the diversity of individual communities and the diversity of sets of different communities. Specifically, if a super-collection of individuals is formed by pooling the individuals of several distinct smaller collections, the idea is that the diversity of the super-collection must be at least as great as the average diversity of the smaller collections. Applied to biological communities,

---

[19] There *can* be reasons to treat individuals of different species differently. Individuals of rare species (and their proportional abundances) are usually weighed more significantly in assessing the diversity of communities in conservation biology, for instance. Rare species are typically more likely to go extinct. Indices that accord changes in their proportional abundances more import than changes in common species are favored in conservation contexts because changes in the former are more likely to influence species persistence than changes in the latter (see §§3-5).

[20] Lewontin [1972] may have been the first to formulate this as an adequacy condition for the concept of diversity.

this requires the diversity of the super-community $\bigcup_{i=1}^{z}\{C_i\}$ formed by pooling the individuals of each community C_i to be greater than or equal to the weighted mean diversity of the C_i (z is an index of the communities). Stated formally for the case of equal weights, and using '*DIV*' to represent diversity:

$$(3)\quad DIV(\bigcup_{i=1}^{z}\{C_i\}) \geq \frac{1}{n}\sum_{i=1}^{z} DIV(C_i).$$

Equality holds only if the C_i are *compositionally* identical, i.e. $V_p^{C_i} = V_p^{C_j}$ for all $i \neq j$.

Lewontin did not provide a rationale for this constraint on ecological diversity, and there are reasons to reject it as an adequacy condition. Consider two simple communities C and D composed of four different species (two each) with *absolute* (not proportional) abundances $\langle 2, 2\rangle$ and $\langle 1000, 1000\rangle$. The absolute species abundance vector for the super community $C \cup D$ with species richness four is $\langle 1000, 1000, 2, 2\rangle$. Equation (3) requires the diversity of $C \cup D$ be greater than the average diversity of C and D, but it is unclear why this is defensible as an adequacy condition on the concept of diversity. $C \cup D$ contains more species than either C or D, and in this respect seems more diverse. But it is also highly uneven compared with C or D. The proportional species abundance vector for $C \cup D$ is approximately $\langle 0.499, 0.499, 0.001, 0.001\rangle$ which is a highly uneven distribution, unlike the highly even distribution of C and D, $\langle 0.5, 0.5\rangle$. Equation (3) therefore forces a strong rank order of species richness over evenness in assessments of diversity.

This *may* be a defensible property of a proposed diversity index.[21] In conservation biology, for example, there may be advantages to prioritizing species richness over evenness in assessments of the diversity of communities targeted for conservation. Equation (3) is not, however, a defensible constraint on *any* potential quantitative specification of diversity. Species richness and evenness are independent properties. Though this does not entail one is not more important than another in evaluations of a community's diversity, nothing about the pre-theoretic concept of ecological diversity seems to suggest otherwise. Pielou [1977, p. 292], for instance, explicitly rejected the constraint imposed by (3): "since diversity depends on two independent properties of a collection ... a collection with few species and high evenness could have the same diversity as another collection with many species and low evenness."[22]

---

[21] Note that this property is consistent with (A1)–(A7) from above, which do not compel any relationship between species richness and evenness in assessments of diversity.

[22] In passing, Lande [1996, p. 8] motivates this condition by pointing out that its denial, "implies the possibility of a negative diversity among communities." But why this is problematic is unclear. It does not, for instance, entail the diversity of any *individual* community is negative, which would clearly be problematic.

3 SIMPSON'S INDEX

The first index that included species richness and evenness as components of diversity found in the ecological literature was proposed by Simpson [1949]. Simpson claimed that the probability two individuals drawn at random (with replacement) from an indefinitely large collection are of the same group is $\sum_{i=1}^{n} p_i^2$, where n is the number of groups exhibited within the collection, and he called it a "measure of concentration." Applied to biological communities, $\sum_{i=1}^{n} p_i^2$ then measures the dominance (in terms of abundance) of species within the community (Pielou 1977) and is at its minimal value $(\frac{1}{n})$ for a given species richness n when individuals of the community are equally distributed among the n species, *i.e.* when $V_p = V_p^{\max}$.

On Simpson's interpretation, the complement of the concentration measure:

$$(D) \qquad 1 - \sum_{i=1}^{n} p_i^2,^{23}$$

represents the probability two randomly selected individuals belong to different species, which is an intuitive measure of diversity.[24] D is at its maximal value for a given species richness n $(D = 1 - \frac{1}{n})$ when individuals are maximally equally distributed among species, i.e. when $V_p = V_p^{\max}$, and at its minimal value when individuals are maximally unequally distributed among species, i.e. when $V_p = V_p^{\min}$.[25]

Several ecologists have suggested other interpretations of D. Hurlbert [1971], for instance, claimed that D multiplied by $\frac{N}{N-1}$ represents the probability of interspecific encounter in the community, rather than just the probability two randomly selected individuals belong to different species. Patil and Taillie [1982a] made a similar claim and showed how quantities such as the waiting time for intra and interspecific encounter are related to D on this stronger interpretation. Recently, Ricotta [2000, p. 246] has suggested the same interpretation. These interpretations are only sound, however, if an additional assumption is made about community structure. In response to Patil and Taillie's analysis, Sugihara [1982] correctly pointed out that D represents the probability of interspecific interaction only if

---

[24] The counterpart of D that does not make the idealization that the two individuals are drawn from an indefinitely large population is $1 - \sum_{i=1}^{n} \frac{N_i(N_i-1)}{N(N-1)}$ and will not be discussed here [Pielou, 1977; Magurran, 2004].

[25] This minimum is not the value of D, however, when all individuals of the community are of the same species, which is 0. $D = 0$ when all but one p_i is zero, but this set of proportional abundances violates the biological requirement on V_p that $p_i \geq \frac{1}{N}$ for all i stipulated above. See the discussion preceding (A4) in Section 2.

[25] The inverse of Simpson's concentration measure, $\frac{1}{\sum_{i=1}^{n} p_i^2}$, is also commonly used as an index of diversity [Williams, 1964; Levins, 1968; Hurlbert, 1971; MacArthur, 1972; Hill, 1973; May, 1975; Pielou, 1977; Magurran, 1988; 2004; Lande, 1996].

the frequencies of interspecific encounters are directly proportional to the relative abundances of the species interacting. The problem, Sugihara [1982, p. 565] emphasized, is that:

> None of these interpretations [of D], however, has yet proved to be very fruitful, as they suffer from such real-world concerns as spatial patchiness and clumping in species distributions, differential mobility, and problems associated with interpreting niche overlap between species from their spatial covariance. Approaching the study of species diversity through a priori models is a valid enterprise, but requires a clear intuition of how communities operate, which thus far seems to be lacking.

Hurlbert's modification of D *may* be a good *estimator* of the probability of interspecific encounter if species are spatially distributed relatively uniformly throughout the area occupied by the community. If distributed in this way, species are likely to interact in direct proportion to their abundances. But without knowing this, and without a thorough understanding of the factors Sugihara mentioned, the legitimacy of the stronger interpretation cannot be reliably verified.[26]

As has been noted throughout the ecological literature, D satisfies (A1)–(A4) [Hill, 1973; Pielou, 1975; 1977; Magurran, 1988; 2004; Lande, 1996]. D also clearly satisfies (A5). As the similar mathematical structure of D and (2) from above suggest, D also satisfies (A6). To see this, recall that for a given species richness (A6) requires that: (i) the diversity of the community represented by V_p be inversely related to $d(V_p, V_p^{\max})$; and, (ii) that if $d(V_p^i, V_p^{\max}) = d(V_p^j, V_p^{\max})$, the diversity of communities represented by V_p^i and V_p^j are equal. Diversity is directly related to $-d(V_p, V_p^{\max})$ given (i), so what is first needed is to show that D exhibits the same relationship with $-d(V_p, V_p^{\max})$. The following algebraic identities demonstrate the required relationship:

$$\begin{aligned} d(V_p, V_p^{\max}) &= \sqrt{\sum_{i=1}^{n} \left(p_i - \tfrac{1}{n}\right)^2} \\ &= \sqrt{\sum_{i=1}^{n} \left(p_i^2 - \tfrac{2p_i}{n} + \tfrac{1}{n^2}\right)} \\ &= \sqrt{\sum_{i=1}^{n} (p_i^2) - \tfrac{2}{n}\sum_{i=1}^{n} (p_i) + \sum_{i=1}^{n} \left(\tfrac{1}{n^2}\right)} \\ &= \sqrt{\sum_{i=1}^{n} (p_i^2) - \tfrac{2}{n} + \tfrac{1}{n}} \\ &= \sqrt{\sum_{i=1}^{n} (p_i^2) - \tfrac{1}{n}}. \end{aligned}$$

---

[26] Similarly, Simpson's interpretation of D implicitly assumes the probability of selecting two individuals from the same group is directly proportional to their relative abundances, which is also violated under a variety of plausible biological conditions. I owe Sahotra Sarkar for emphasizing this fact.

Substitution for $\sum_{i=1}^{n} p_i^2$ using the definition of D from above yields: $\sqrt{(1-D)-\frac{1}{n}}$, which equals $\sqrt{-D+\frac{n-1}{n}}$.[27] Since $\frac{n-1}{n}$ is constant and nonnegative for a given n, $d(V_p, V_p^{\max})$ decreases as D increases, establishing the required relationship (i). Similarly, if $d(V_p^i, V_p^{\max}) = d(V_p^j, V_p^{\max})$, then $\sqrt{-D_i+\frac{n-1}{n}} = \sqrt{-D_j+\frac{n-1}{n}}$ where D_i and D_j represent the complement of Simpson's index for the proportional abundances of V_p^i and V_p^j, respectively. From this it follows that $D_i = D_j$ (given that D must be positive), establishing (ii). (A6) is thereby satisfied.[28]

That $d(V_p, V_p^{\max}) = \sqrt{-D+\frac{n-1}{n}}$ also shows, however, that D does not satisfy (A7). (A7) requires diversity decrease linearly with $d(V_p, V_p^{\max})$, but D scales quadratically with $d(V_p, V_p^{\max})$. Thus, D violates (A7) because it is more sensitive to changes in proportional species abundance vectors at greater distances from $V_p^{\max}$. In particular, D is more sensitive to the proportional abundances of especially abundant or rare species. If diversity were specified with Simpson's index, (A7) would therefore not be satisfied and diversity would fail to be impartial among individual organisms comprising a community.

4 SHANNON'S INDEX

Probably the most popular index of community diversity is Shannon's index (H), so-called after Claude Shannon, who developed it in the context of what came to be called information theory.[29] The index was originally formulated to quantify the amount of information transmitted in a communication channel [Shannon, 1948; Shannon and Weaver, 1949] and, despite Margalef's [1958] claim to priority, Good [1953] was the first to use it as an index of ecological diversity. The index:

$$(H) \qquad -\sum_{i=1}^{n} p_i \ln p_i,^{30}$$

can be used to measure the information of a message composed of n types of symbols whose individual probability of occurrence is $p_i, i = 1, ..., n$. Within ecology, however, p_i represents the familiar proportional abundance of species i and n represents the community's species richness.

---

[27] The term $\sqrt{-D+\frac{n-1}{n}}$ is well defined because $-D+\frac{n-1}{n}$ is nonnegative. Specifically, at its maximum $D = 1-\frac{1}{n}$, in which case $\sqrt{-D+\frac{n-1}{n}} = 0$.

[28] As a check on this result, it can be verified that the two species abundance vectors $\langle \frac{8}{15}, \frac{4}{15}, \frac{3}{15} \rangle$ and $\langle \frac{7}{15}, \frac{6}{15}, \frac{2}{15} \rangle$ from above, which have the same distance (0.3887) according to equation (2), also have the same value (0.6044) according to Simpson's index.

[29] Aczél and Daróczy [1975] suggest Norbert Wiener independently developed an index which is a special case of Shannon's more general index in 1948 [Wiener, 1948].

[30] Shannon's index assumes p_i is a proportional species abundance from an infinitely large community [Magurran, 2004]. This idealization, and the emendation needed to correct it for

H is at its maximal value ($\ln n$) for a given species richness n when individual organisms are equally distributed among species, thereby satisfying adequacy condition (A3) [Pielou, 1977]. Similar to D, that H satisfies (A1), (A2), and (A4) is well-known [Hill, 1973; Pielou, 1975; 1977; Magurran, 1988; 2004; Lande, 1996]. H obviously satisfies (A5). But, as a simple example demonstrates, Shannon's index is more sensitive than Simpson's index to the abundances of rare species [Peet, 1974] and therefore fails (A6).[31] Consider a four species community composed of one abundant, one rare, and two evenly distributed species such that $V_p = \langle \frac{12}{32}, \frac{8}{32}, \frac{8}{32}, \frac{4}{32} \rangle$. $D = 0.7188$ and $H = 1.3209$. If the rare species becomes rarer, so that V_p becomes $V_p^r = \langle \frac{13}{32}, \frac{9}{32}, \frac{9}{32}, \frac{1}{32} \rangle$, then evenness decreases and both D and H decrease to 0.6758 and 1.1878, respectively. If the abundant species becomes more abundant, so that V_p becomes $V_p^a = \langle \frac{15}{32}, \frac{7}{32}, \frac{7}{32}, \frac{3}{32} \rangle$, then evenness also decreases and both D and H decrease to 0.6758 and 1.2420, respectively. That H decreases more than D for V_p^r vs. V_p^a does not necessarily show it is more sensitive to rare or abundant species because the range of values D and H take between $V_p^{\min}$ and $V_p^{\max}$ differ. Differences in their values may be due merely to a scaling effect. But the way values of D and H change from V_p to $V_p^{\min}$ and to $V_p^{\max}$ reveal that H is more sensitive to abundances of rare species. Specifically, H decreases more between V_p and V_p^r (1.3209–1.1878 = 0.1331) than between V_p and V_p^a (1.3209–1.2420 = 0.0789), while D decreases by the same amount. H is therefore more sensitive to the abundances of rare species, unlike D, and contrary to (A6)(i). In addition, since $d(V_p^r, V_p^{\max}) = d(V_p^a, V_p^{\max})$, (A6)(ii) requires a diversity index assign the same value for V_p^r and V_p^a. Their H values, however, differ (1.1878 and 1.3209, respectively).[32] (A6) is a necessary condition of (A7), so H also fails to satisfy (A7).

How should H be *ecologically* interpreted? Pielou [1977] provided a particularly simple and probably the clearest interpretation. As an entropy measure, Pielou suggested H measures uncertainty, and that diversity and uncertainty are closely related concepts.[33] Specifically, as the diversity of a community increases,

---

finite communities, will not be discussed here. The sampling estimator of H which does not make this assumption is Brillouin's index, $\frac{1}{N} \ln \left(\frac{N!}{N_1!N_2!\ldots N_n!} \right)$ [Brillouin, 1962].

[31] Compared with D, H is also less sensitive to proportional abundances of species (and hence evenness) and more sensitive to species richness [May, 1975; Magurran, 1988; 2004].

[32] Despite these and other differences (see below), Simpson's and Shannon's indices are both members of a family of entropy measures defined by: $H_q = \frac{\ln \sum_{i=1}^{n} p_i^{\alpha}}{1-\alpha}$, in which $\alpha > 0$ and $\alpha \neq 1$ [Rényi, 1961; Pielou, 1975]. Rényi [1961] showed, for example, that $\lim_{\alpha \to 1} (H_q) = -\sum_{i=1}^{n} p_i \ln p_i = H$. Pielou [1975] showed that if $\alpha = 2$, $H_q = -\ln \sum_{i=1}^{n} p_i^2$, which is equivalent to the inverse form of Simpson's diversity index after exponential transform.

[33] Pielou's view of the proper ecological interpretation of Shannon's index shifted markedly in the 1960s and 1970s. In 1966, Pielou suggested there was an "obvious analogy" between a biological community and a coded message, and that, "the actions of a biologist are formally identical with those of a man observing, one after another, the symbols of a message" [1966, p. 164]. (A very similar characterization had been suggested by Margalef [1958] almost a

the uncertainty about which species a randomly selected individual belongs to increases. It is difficult to deny this claim, but the crucial issue is whether H is the *uniquely* appropriate measure of uncertainty. After all, given the interpretation of D described earlier, it also seems to measure a similar, if not identical, kind of uncertainty about a biological community.

Pielou's argument that H is the uniquely appropriate index of ecological diversity relied on a mathematical fact about H proved in a non-ecological context by Shannon. Shannon [1948] showed that H is the only function (up to a multiplicative constant) which exhibits three properties he thought reasonable to require of the concept of information.[34] These properties include that an information function should be continuous in p_i, and that it should monotonically increase with n for maximally even p_i (see [A1′] from above). The most important was an additive property which was the basis of Pielou's argument for preferring H over other diversity indices and, even stronger, that it constitutes an adequacy condition for any index of diversity [Pielou, 1977, p. 293-294]. The interpretation of this property as applied to biological communities requires some elaboration.

Just as V_p is based on a classification P of individuals of a biological community into n species, let V_q be a proportional abundance vector with m components based on a different classification Q. The second classification could be derived from further taxonomic information, information on habitat requirements or other properties of individual organisms, *etc.* As with the p_i, assume that the proportional abundances q_j of the second classification are such that $\sum_{j=1}^{m} q_j = 1$, and that each individual organism falls into only one class. In analogy with the information concept, if the two classifications are independent, Pielou [1977, p. 294] suggested that ecological diversity must satisfy:

$$(4) \quad DIV(PQ) = DIV(P) + DIV(Q);$$

in which $DIV(PQ)$ is the diversity of a biological community with individuals classified into both P and Q for a total of $m \times n$ classes, and $DIV(P)$ and $DIV(Q)$ are the diversity of the community with individuals classified by only P or Q, respectively.[35] Mathematically, it can be shown that H is the only continuous function of the p_i up to a multiplicative constant that satisfies (A1′), (A3), and for which equation (4) holds [Khinchin, 1957].

In general, (4) is a desirable property because it permits the additive decomposition of a function of two combined input arguments. Together with its intuitive

---

decade before.) By 1975, however, Pielou's view of such analogies was decidedly negative: "it cannot be too strongly emphasized that fancied links between the information-theoretic concept of 'information' and the diversity of an ecological community are merely fancies and nothing more" [1975, p. 9].

[34] In a later review of information theory, Aczél and Daróczy [1975, p. 29] called these properties, "natural properties which are essential from the point of view of information theory." Shannon, however, only called them "reasonable" and emphasized that, "The real justification of these definitions, however, will reside in their implications," [Shannon and Weaver, 1949, p. 50].

[35] Equation (4) can be generalized to any finite number of independent classifications.

plausibility for the concept of information, this motivated Shannon [1948], Khinshin [1957], Rényi [1961], Aczél and Daróczy [1975], and others to stipulate it as an adequacy condition for any information measure. What is needed to show it is an appropriate adequacy condition for the concept of ecological diversity, however, is an account of why this is a *necessary* property of diversity. Pielou [1977] did not supply such a rationale. *If* diversity were specified as H, (4) logically follows given its logarithmic form, and it can be agreed that there are benefits of being able to additively decompose ecological diversity in this way [Pielou, 1977, pp. 303-307]. This is insufficient, however, to establish the stronger claim that (4) is a defensible *adequacy* condition, especially given that there are other non-additive methods with attractive features in which ecological properties besides proportional species abundances can be integrated into a measure of community diversity (e.g. [Rao, 1982; Ricotta, 2002]).

5 THE ROLE OF THE DIVERSITY CONCEPT WITHIN ECOLOGY

Simpson's and Shannon's diversity indices emerged within ecology in the 1950s. By the late 1960s, a large number of diversity indices had been formulated, and numerous empirical studies of different ecological systems were being conducted to estimate diversity using these indices.[36] Table 1 lists some of these indices. One reaction to this diversity of diversity indices was the thought that anything goes, that the diversity concept was deeply and problematically unclear. The attention being devoted to indices of diversity sparked several such criticisms, perhaps the most incisive from Hurlbert [1971] who called diversity a "nonconcept." His influential critique of this research agenda targeted the fundamental vagueness of the underlying concept, which he thought ecologists had exacerbated by appropriating statistical measures of diversity developed in nonbiological contexts, such as information theory, with dubious ecological relevance. MacArthur [1972, p. 197] voiced the same criticism around the same time: "Applying a formula and calculating a 'species diversity' from a census does not reveal very much; only by relating this diversity to something else – something about the environment perhaps – does it become science." Other ecologists were similarly skeptical of the role of the diversity concept within ecology, and generally of the ecological utility of information theory (e.g. [Hill, 1973]).

Rather than attempt to rehabilitate the concept by proposing adequacy conditions by which to evaluate relative merits and weaknesses of different diversity indices as attempted above, Hurlbert suggested the search for relationships between diversity and other community properties, such as stability, should be refocused on the relationship between those properties and indices that reflect biologically meaningful properties that might influence community dynamics.[37] His proposed index of the probability of interspecific encounter is one example (see §3). As

---

[36] See [Pielou, 1975; 1977; Magurran, 1988; 2004; Sarkar, 2007; Drake, 2007] for reviews.

[37] See [Justus, 2008] for an analysis of the stability-diversity debate within ecology.

a measure of ecological diversity, note that species richness alone fails this test since it is generally unlikely that extremely rare species (*e.g.* s_1 in community A from §2) play an important role in community dynamics.[38] Since it does not consider evenness, furthermore, species richness fares very poorly as a specification of ecological diversity; it fails adequacy conditions (A2)–(A7). Species richness was and remains, however, the predominant surrogate for diversity in analyses of stability-diversity relationships (e.g. [Tilman, 1996; 1999]). As such, it is unclear these studies offer significant insights into possible relationships between stability and diversity in biological communities.[39]

Since Hurlbert's critique, ecologists have proposed a multitude of new diversity indices to satisfy different proposed adequacy conditions besides those about species richness and evenness (see [Magurran, 2004; Ricotta, 2005; Sarkar, 2007] for reviews). Diversity indices should increase, for instance, as interspecific taxonomic and functional differences increase. Besides properties of species, spatial properties of their geographical distribution could also be included in an index of ecological diversity. Since species distributions are significantly influenced by regional geology and environmental gradients, however, including these properties would expand the scope of ecological diversity beyond just the biological properties of communities. Expanded in this way, the "diversity" of the physical environment in which a community resided would also contribute to the value of indices of *ecological* diversity.[40]

How compatible these additional adequacy conditions are with one another, or with the other conditions is not yet clear. Some conditions appear to be conceptually independent, but some formal diversity indices suggest that others are not. Rao's [1982] "quadratic entropy" diversity index, for instance, which generalizes the Simpson index [Ricotta and Avena, 2003], incorporates interspecific taxonomic and functional differences as well as evenness and species richness into a single quantitative measure. Unlike the Shannon and Simpson indices, however, quadratic entropy violates the adequacy condition (A3) [Ricotta, 2005]. (A4) is also violated. This is as it should be. If functional or taxonomic information is included in assessments of diversity, then high functional or taxonomic diversity may make a less even community more diverse overall than a more even one. In effect, functional or taxonomic diversity can trump evenness.

As new indices are devised, similar incompatibilities between other adequacy conditions may be revealed. Absent a general proof that plausible adequacy criteria are themselves incompatible, however, the formulation of a uniquely defensible

---

[38] Potentially important exceptions include keystone species and so-called "ecosystem engineers" (see [Paine, 1969; Jones *et al.*, 1994; Power *et al.*, 1996]).

[39] A recent study by Tilman *et al.* [2006] that uses the Shannon index to measure the diversity of a Minnesota grassland and finds the same positive correlation as between species richness and temporal stability is a rare exception.

[40] Properties of the spatial extent of species, in particular their geographical rarity, are clearly relevant to the concept of diversity utilized in the context of biodiversity conservation, and therefore must be integrated into any defensible measure of biodiversity [Sarkar, 2002; 2005; 2007].

diversity index satisfying all of them remains possible. Focusing on species richness and evenness, adequacy conditions (A1)–(A7) are compatible and thus a defensible diversity measure satisfying them exists. In fact, (A6) and (A7) suggest two obvious candidates: (i) $-\sqrt{\sum_{i=1}^{n}\left(p_i - \frac{1}{n}\right)^2}$; and, (ii) $\frac{1}{\sqrt{\sum_{i=1}^{n}\left(p_i - \frac{1}{n}\right)^2}}$, which both also satisfy (A1)–(A5). Despite the multitude of quantitative measures of ecological diversity in the literature, that a set of defensible adequacy conditions can be formulated shows the concept is not problematically obscure and that it may facilitate scientific insights into biological communities.

Glossary:

n = species richness;
p_i = the proportional abundance of species i;
N_i = the abundance of species i;
N = the abundance of all species in the community;
N_{max} = abundance of the most abundant species;
d_{ij} = distance (*e.g.* taxonomic, functional, *etc.*) between species i and j;
r_i = rank of species i in V_p;
a = a constant ≥ 0.

Table 1. A list of some of the diversity indices in the ecological literature.

Diversity Index	Mathematical Operationalization	Origin
Simpson's Index (infinite community)	$D\colon\ 1-\sum\limits_{i=1}^{n} p_i^2$ or $\frac{1}{\sum\limits_{i=1}^{n} p_i^2}$	**Simpson [1949]**
Simpson's Index (finite community)	$1-\sum\limits_{i=1}^{n} \frac{N_i(N_i-1)}{N(N-1)}$ or $\frac{1}{\sum\limits_{i=1}^{n} \frac{N_i(N_i-1)}{N(N-1)}}$	**Simpson [1949]**
Shannon's index (infinite community)	$H\colon\ -\sum\limits_{i=1}^{n} p_i \ln p_i$	**Shannon [1948]**
Brillouin's index (Shannon's index for a finite community)	$\frac{1}{N}\ln\left(\frac{N!}{N_1!N_2!\ldots N_n!}\right)$	**Brillouin [1962]**
Margalef's Index	$\frac{n-1}{\ln(N)}$	**Margalef [1958]**
Menhinick's Index	$\frac{n}{\sqrt{N}}$	**Menhinick [1964]**
McIntosh's index	$\frac{N-\sqrt{\sum\limits_{i=1}^{n} N_i}}{N-\sqrt{N}}$	**McIntosh [1967]**
Berger-Parker's index	$1-\frac{N_{\max}}{N}$ or $\frac{N}{N_{\max}}$	**Berger and Parker [1970]**
Hurlbert's "Interspecific Encounter" Index	$\left(\frac{N}{(N-1)}\right)\left(1-\sum\limits_{i=1}^{n} p_i^2\right)$	**Hurlbert [1971]**
Fager's Index	$\frac{N(n+1)}{2}-\sum\limits_{i=1}^{n} r_i N_i$	**Fager [1972]**
Hill's Family of Diversity Indices	$\left(\sum\limits_{i=1}^{n} p_i^a\right)^{\frac{1}{(1-a)}}$	**Hill [1973]**
Keefe and Bergersen's Index	$1-\frac{N}{(N-1)}\left(\sum\limits_{i=1}^{n} p_i^2-\frac{1}{N}\right)$	**Keefe and Bergersen [1977]**
Rao's "Quadratic Entropy" Index	$\boldsymbol{Q}\colon\ \sum\limits_{i,j=1}^{n} d_{ij} p_i p_j$	**Rao [1982]**

ACKNOWLEDGEMENTS

Thanks to Russell Lande, Samir Okasha, Carl Salk, and Sahotra Sarkar for helpful comments. Members of the Florida State University biology department and Sydney-ANU philosophy of biology workshop also provided helpful feedback on a presentation based on this work.

BIBLIOGRAPHY

[Aczél and Daróczy, 1975] J. Aczél, and Z. Daróczy. *On Measures of Information and Their Characteristics.* Academic Press, New York, 1975.

[Berger and Parker, 1970] W. H. Berger and F. L. Parker. Diversity of Planktonic *Formaminifera* in Deep-sea Sediments. *Science* **168**: 1345-1347, 1970.

[Brillouin, 1962] L. Brillouin. *Science and Information Theory.* 2^{nd} edition. Academic Press, New York, 1962.

[Drake, 2007] J. M. Drake. The Background of Biodiversity: Quantitative Measures of Species Diversity, 1943-1982. Unpublished Manuscript.

[Dyer and Coley, 2001] L. A. Dyer and P. D. Coley. Latitudinal Gradients in Tri-trophic Interactions." In Tscharntke, T. and Hawkins, B. A. (eds.) *Multitrophic Level Interactions.* New York, NY: Cambridge University Press, pp. 67-88, 2001.

[Eberhardt, 1969] L. L. Eberhardt. Some Aspects of Species Diversity Models. *Ecology* **50**: 503-505, 1969.

[Fager, 1973] E. W. Fager. Diversity: A Sampling Study. *American Naturalist* **106**: 293-310, 1973.

[Fisher *et al.*, 1943] R. A. Fisher; A. S. Corbet; and C. B. Williams. The Relationship Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. *Journal of Animal Ecology* **12**: 42-58, 1943.

[Good, 1953] I. J. Good. The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika* **40**: 237-264, 1953.

[Hill, 1973] M. O. Hill. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology* **54**: 427-432, 1973.

[Hurlbert, 1971] S. Hurlbert. The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology* **52**: 577-586, 1971.

[Jones *et al.*, 1994] C. G. Jones; J. H. Lawton; M. Shachak. Organisms as Ecosystem Engineers. *Oikos* **69**: 373-386, 1994.

[Justus, 2006] J. Justus. Loop analysis and Qualitative Modeling: Limitations and Merits. *Biology and Philosophy* **21**: 647-666, 2006.

[Justus, 2008] J. Justus. Complexity, Diversity, and Stability. In Sarkar, Sahotra and Plutynski, Anya (eds.) A Companion to the Philosophy of Biology. Malden, MA: Blackwell, pp. 321–350, 2008.

[Kaplansky, 1977] I. Kaplansky. *Set Theory and Metric Spaces.* New York: Chelsea Publishing, 1977.

[Keefe and Bergersen, 1977] T. J. Keefe and E. P. Bergersen. A Simple Diversity Index Based on the Theory of Runs. *Water Research* **11**: 689-691, 1977.

[Khinchin, 1957] A. I. Khinchin. *Mathematical Foundations of Information Theory.* New York, NY: Dover Publications, 1957.

[Lande, 1996] R. Lande. Statistics and Partitioning of Species Diversity, and Similarity among Multiple Communities. *Oikos* **76**: 5-13, 1996.

[Lewontin, 1972] R. Lewontin. The Apportionment of Human Diversity. *Evolutionary Biology* **6**: 381-398, 1972.

[Lloyd and Ghelardi, 1964] M. Lloyd and R. Ghelardi. A Table for Calculating the Equitability of Species Diversity. *Journal of Animal Ecology* **33**: 217-225, 1964.

[MacArthur, 1957] R. H. MacArthur. On the Relative Abundance of Bird Species. *Proceedings of the National Academy of Sciences of the USA* **43**: 293-295, 1957.

[MacArthur, 1972] R. H. MacArthur. *Geographical Ecology.* Harper and Row, New York, 1972.

[Maclaurin and Sterelny, 2008] J. Maclaurin and K. Sterelny. *What is Biodiversity?* Chicago: The University of Chicago Press, 2008.

[Magurran, 1988] A. E. Magurran. *Ecological Diversity and Its Measurement.* Princeton University Press, Princeton, 1988.

[Magurran, 2004] A. E. Magurran. *Measuring Biological Diversity.* Blackwell Publishing, Malden, 2004.

[Margalef, 1958] R. Margalef. Information Theory in Ecology. *General Systems* **3**: 36-71, 1958.

[May, 1975] R. M. May. Patterns of Species Abundance and Diversity. In Cody, Michael L. and Diamond, Jared M. (eds.) *Ecology and Evolution of Communities.* Cambridge, MA: Harvard University Press, pp. 81-120, 1975.

[McIntosh, 1967] R. P. McIntosh. An Index of Diversity and the Relations of Certain Concepts to Diversity. *Ecology* **48**: 392-404, 1967.

[Menhinick, 1964] E. P. Menhinick. A Comparison of some Species-Individuals Diversity Indices Applied to Samples of Field Insects. *Ecology* **45**: 859-861, 1964.

[Møller, 1998] A. P. Møller. Evidence of Larger Impact of Parasites on Hosts in the Tropics: Investment in Immune Function Within and Outside the Tropics? *Oikos* **82**: 265-270, 1998.

[Oksanen *et al.*, 1981] L. Oksanen; S. Fretwell; J. Arrunda; and P. Niemela. Exploitation Ecosystems in Gradients of Primary Productivity. *American Naturalist* **118**: 240-261, 1981.

[Paine, 1969] R. T. Paine. A Note on Trophic Complexity and Species Diversity. *American Naturalist* **100**: 91-93, 1969.

[Patil and Taillie, 1982a] G. P. Patil and C. Taillie. Diversity as a Concept and Its Measurement. *Journal of the American Statistical Association* **77**: 548-561, 1982.

[Patil and Taillie, 1982b] G. P. Patil and C. Taillie. Diversity as a Concept and Its Measurement: Rejoinder. *Journal of the American Statistical Association* **77**: 548-561, 1982.

[Peet, 1974] R. K. Peet. The Measurement of Species Diversity. *Annual Review of Ecology and Systematics* **5**: 285-307, 1974.

[Peet, 1975] R. K. Peet. Relative Diversity Indices. *Ecology* **56**: 496-498, 1975.

[Pianka, 1966] E. Pianka. Latitudinal Gradients in Species Diversity: A Review of Concepts. *American Naturalist* **100**: 33-46, 1966.

[Pielou, 1966] E. C. Pielou. The Use of Information Theory in the Study of the Diversity of Biological Populations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* **4**: 163-177, 1966.

[Pielou, 1975] E. C. Pielou. *Ecological Diversity.* John Wiley & Sons, New York, 1975.

[Pielou, 1977] E. C. Pielou. *Mathematical Ecology.* 2^{nd} Edition. John Wiley & Sons, New York, 1977.

[Preston, 1948] F. W. Preston. The Commonness and Rarity of Species. *Ecology* **29**: 254-283, 1948.

[Preston, 1962a] F. W. Preston. The Canonical Distribution of Commonness and Rarity, Part I. *Ecology* **43**: 185-215, 1962.

[Preston, 1962b] F. W. Preston. The Canonical Distribution of Commonness and Rarity, Part II. *Ecology* **43**: 410-432, 1962.

[Rao, 1982] C. R. Rao. Diversity and Dissimilarity Coefficients – A Unified Approach. *Theoretical Population Biology* **21**: 24-43, 1982.

[Rényi, 1961] A. Rényi. On Measures of Entropy and Information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics, and Probability* **1**: 547-561, 1961.

[Ricotta, 2000] C. Ricotta. From Theoretical Ecology to Statistical Physics and Back: Self-Similar Landscape Metrics as a Synthesis of Ecological Diversity and Geometrical Complexity. *Ecological Modeling* **125**: 245-253, 2000.

[Ricotta, 2002] C. Ricotta. Bridging the Gap between Ecological Diversity Indices and Measures of Biodiversity with Shannon's Entropy: Comment to Izsák and Papp. *Ecological Modeling* **152**: 1-3, 2002.

[Ricotta, 2005] C. Ricotta. Through the Jungle of Biological Diversity. *Acta Biotheoretica* **53**: 29-38, 2005.

[Ricotta and Avena, 2003] C. Ricotta and G. C. Avena. An Information-Theoretical Measure of Taxonomic Diversity. *Acta Biotheoretica* **51**: 35-41, 2003.

[Sarkar, 2002] S. Sarkar. Defining "Biodiversity"; Assessing Biodiversity. *The Monist* **85**: 131-155, 2002.

[Sarkar, 2005] S. Sarkar. *Biodiversity and Environmental Philosophy.* Cambridge University Press, New York, 2005.

[Sarkar, 2007] S. Sarkar. From Ecological Diversity to Biodiversity. In Hull, David and Ruse, Michael, *Cambridge Companion to the Philosophy of Biology.* Cambridge, MA: Cambridge University Press, 2007.

[Shannon, 1948] C. E. Shannon. A Mathematical Theory of Communication. *Bell Systems Technical Journal* **27**: 379-423, 623-656, 1948.

[Shannon and Weaver, 1949] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, 1949.

[Simpson, 1949] E. H. Simpson. Measurement of Diversity. *Nature* **163**: 688, 1949.

[Smith *et al.*, 1996] B. Smith and J. B. Wilson. A Consumer's Guide to Evenness Indices. *Oikos* **76**: 70-82, 1996.

[Sugihara, 1982] G. Sugihara. Comment. *Journal of the American Statistical Association* **77**: 564-565, 1982.

[Tilman, 1996] D. Tilman. Biodiversity: Population Versus Ecosystem Stability. *Ecology* **77**: 350 -363, 1996.

[Tilman, 1999] D. Tilman. The Ecological Consequences of Biodiversity: A Search for General Principles. *Ecology* **80**: 1455-1474, 1999.

[Tilman *et al.*, 2006] D. Tilman; P. B. Reich; J. M. H. Knops. Biodiversity and Ecosystem Stability in a Decade-long Grassland Experiment. *Nature* **441**: 629-632, 2006.

[Tramer, 1969] E. J. Tramer. Bird Species Diversity: Components of Shannon's Formula. *Ecology* **50**: 927-929, 1969.

[Wiener, 1948] N. Wiener. *Cybernetics: Or the Control and Communication in the Animal and the Machine.* MIT Press, Cambridge, 1948.

[Willig *et al.*, 2003] M. R. Willig; D. M. Kaufmann; and R. D. Stevens. Latitudinal Gradients of Biodiversity: Pattern, Process, Scale and Synthesis. *Annual Review of Ecology and Systematics* **34**: 273-309, 2003.