

## THE MEASUREMENT OF NICHE OVERLAP AND SOME RELATIVES<sup>1</sup>

STUART H. HURLBERT  
 Department of Biology, San Diego State University  
 San Diego, California 92182 USA

*Abstract.* Existing overlap indices are examined and are judged to be inadequate on the grounds (1) that they lack simple and appropriate biological interpretations, and (2) that they ignore possible variation among resource states in abundance or availability ( $a_i$ ). Two indices interpretable in terms of encounters are proposed. Niche overlap ( $L$ ) is measured as the degree to which frequency of interspecific encounter is higher or lower than it would be if each species utilized each resource state in proportion to its abundance ( $a_i$ ). Directional overlap ( $Z_{xy}$ ) is measured as the density of species  $Y$  encountered, on the average, by an individual of species  $X$ . When resource states are equal in size,  $L$  is equivalent to Lloyd's 'interspecies patchiness,' and  $Z_{x(y)}$  is equivalent to his 'mean crowding on species 1 by species 2.' Indices which incorporate variation in resource state abundance are also developed for mean crowding, patchiness and niche breadth.

*Key words:* Association; niche breadth; niche overlap; patchiness.

### INTRODUCTION

In studies of species interactions and community structure it is useful to quantify the degree to which 2 species overlap in their utilization of space, food or other resources. To that end several measures of niche overlap have been proposed. These are briefly discussed below. The purpose of this article is to present an improved and more general index and to discuss certain related measures.

The notation employed here is presented in the first part of Table 1 and requires only brief explanation. "Resource state" is used in a broad sense. The resource states may correspond to any 1 of several classes of entities, such as: a set of quadrats or other arbitrary sampling units; a set of groups of quadrats, where quadrats are grouped on the basis of one or more characteristics (e.g., soil moisture); a set of natural sampling units (e.g., lakes, host plants, etc.), considered individually or grouped on the basis of similarity in some respect; a set of prey species; or, a set of prey categories, defined with respect to taxonomic (supraspecific) group, size, or some other property.

Resource state abundance or availability has correspondingly varied meanings. In discussions of spatial overlap, it usually will be measured as area or volume. In relation to dietary overlap, abundances can be measured as the standing crop of each dietary category, corrected perhaps for differences in productivity or renewal rate, cost of capture, food value, and so on (Schoener 1974).

### EXISTING MEASURES

At least 5 types of niche overlap indices have been proposed; these are listed and later evaluated. The simplest is:

$$C_{xy} = 1 - \frac{1}{2} \left( \sum_i |p_{xi} - p_{yi}| \right), \quad (1)$$

$$= \sum_i \min(p_{xi}, p_{yi}),$$

where  $p_{xi} = x_i/X$   
 and  $p_{yi} = y_i/Y$ ,

which, perhaps because of its simplicity, is re-invented every so many years (Renkonen 1938 *vide* Goodall 1973, Whittaker and Fairbanks 1958, Schoener 1970).

The most widely used overlap indices are those derived from or related to the "competition coefficients" of the Lotka-Volterra equations. These include:

$$O_{xy} = \sum_i (p_{xi}p_{yi}) / \left( \sum_i p_{xi}^2 \cdot \sum_i p_{yi}^2 \right)^{1/2} \quad (2)$$

(Pianka 1973);

$$C_m = 2 \sum_i (p_{xi}p_{yi}) / \left( \sum_i p_{xi}^2 + \sum_i p_{yi}^2 \right) \quad (3)$$

(Morisita 1959, Horn 1966);

$$\text{and } R_{xy} = \frac{\left[ X(X-1) + Y(Y-1) \right] \left[ \sum_i (p_{xi}p_{yi}) - \frac{1}{n} \right]}{\left[ \sum_i x_i(x_i-1) + \sum_i y_i(y_i-1) \right]} \quad (4)$$

when overlap *greater* than expected on the hypothesis of independence, or

$$R_{xy} = n \left( \sum_i p_{xi}p_{yi} \right) - 1, \quad (5)$$

when overlap *less* than expected (Morisita 1971).

In addition, to facilitate theoretical analyses of competition and related phenomena, modifications of the competition coefficient have been developed which assume that resource utilization is normally distributed along the resource axis (MacArthur 1972, May and MacArthur 1972). As these are not intended for use in empirical studies, they will not be discussed here.

<sup>1</sup> Manuscript received 16 November 1976; accepted 8 August, 1977.

Building on his concept of "mean crowding," Lloyd (1967) derived and used an overlap index he called "interspecies patchiness," given as:

$$I_{xy} = n \sum_i (p_{xi} p_{yi}). \quad (6)$$

Equation 6 has attractive properties but has been overlooked by later workers. Its mathematical relationship to Eqs. 2 through 5 is largely coincidental.

Constituting a fourth class of overlap indices are those based on the Shannon-Wiener information function. Horn (1966) proposed

$$R_{ii} = \frac{\sum_i \left\{ [(x_i + y_i)/N] \log [(x_i + y_i)/N] \right\} - \sum_i \left[ (x_i/N) \log (x_i/N) + (y_i/N) \log (y_i/N) \right]}{(X/N) \sum_i (p_{xi} \log p_{xi}) + (Y/N) \sum_i (p_{yi} \log p_{yi}) - \sum_i \left[ (X_i/N) [\log (x_i/N)] + (Y_i/N) \log (y_i/N) \right]}, \quad (7)$$

where  $N = X + Y$ .

Colwell and Futuyma (1971) presented an elaborated version of this (and of Eq. 1) that attempts to minimize distortions arising from the arbitrary manner in which a resource is partitioned into resource states.

A final measure is the product-moment correlation coefficient:

$$C_p = \frac{\left[ \sum_i \left( x_i - \frac{X}{n} \right) \left( y_i - \frac{Y}{n} \right) \right]}{\left[ \sum_i \left( x_i - \frac{X}{n} \right)^2 \sum_i \left( y_i - \frac{Y}{n} \right)^2 \right]^{1/2}}. \quad (8)$$

This has not been used explicitly to measure overlap but it has been widely used in phytosociology as a measure of association: Goodall (1973) suggests that it and Eq. 1 are perhaps the most appropriate measures of "species distributional similarity."

#### CRITERIA AND EVALUATION

In my opinion, the 2 most important features of an index of overlap and the degree to which these are found in existing indices are as follows:

##### *Appropriateness and simplicity*

An appropriate index would seem to be one which defines overlap in a manner consistent with the meaning generally given, explicitly or implicitly, to that term in index-free discussions of resource relationships, competition, and so on. By simplicity, I refer to simplicity of interpretation, not ease of calculation. Not surprisingly, most of the indices have neither an appropriate nor a simple biological interpretation. This is the result of inattention to definition. Workers who have required quantification of niche overlap generally have proceeded directly from an intuitive notion to the selection of an index without making any prior attempt

to formalize their concept of overlap in biological terms.

Equations 2, 3, and 6 have simple biological interpretations as they are expressed in terms of "crowding" or "probability of encounter," which may be considered equivalent concepts. Equation 6, better than the others, meets our criterion of appropriateness: it measures, under certain conditions, the degree to which the probability of interspecific encounter (or degree of interspecific crowding) is higher or lower than would be expected if both species were uniformly distributed over the available space or resources.

Equations 2, 3, and 4 reflect the ratio of the probability of interspecific encounter to the probability of intraspecific encounter or, in other terms (MacArthur 1972, May and MacArthur 1972), the ratio of niche proximity to niche breadth. As such they are inappropriate as indices of overlap. The basis for this conclusion is illustrated by a comparison of cases I and II in Table 1. By any usual connotation of overlap, overlap in resource utilization is equal in the 2 cases. This conflicts with the values yielded by such indices (Table 1, Eq. 3) and certain others (Eq. 8) as well. I can see no justification for the implicit assumption that overlap is partly a function of the species' niche widths outside the overlap zone, i.e., of the species' utilization functions for the nonshared resource states. Why should the frequency of intraspecific encounters or the intensity of intraspecific competition bear any fixed relationship to interspecific overlap in resource use?

The application of the Shannon-Wiener function to the quantification of niche overlap seems as unnecessarily esoteric as is the use of this function as a diversity index (Hurlbert 1971). Equation 7 and its relatives measure "relative mutual information" (Colwell and Futuyma 1971) which, like the "bits" or "nats" in which the numerator and denominator of Eq. 7 are measured, have no clear meaning so far as has been discovered.

Equations 1 and 8 are simple enough, and the first has the appealing geometric property of measuring the area of intersection of the probability density functions of species X and species Y with respect to a given resource axis. However, neither is expressed in units appropriate to our concept of niche overlap, largely, at least in the case of Eq. 1, because they ignore variation in resource state abundance. Thus both yield identical values for cases VI and VII in Table 1. I feel that overlap must be considered greater in case VII than in case VI, since, other things being equal, the 2 species are more likely to encounter or crowd each other or deplete each other's resources in case VII than they are in case VI.

##### *Variation in resource state abundance*

It seems desirable that an overlap index take into account the variation among resource states in abundance or size ( $a_i$ ). With the exception of Schoener's

TABLE 1. Comparison of selected indices of niche overlap, calculated for various distributions of species-pairs over resource states

Symbolic notation		Case I		Case II		Case III		Case IV		Case V		Case VI		Case VII								
Resource state (i)	No. individuals of resource state	sp. X		sp. Y		sp. X		sp. Y		sp. X		sp. Y		sp. X								
		$x_i$	$y_i$	$x_i$	$y_i$	$x_i$	$y_i$	$x_i$	$y_i$	$x_i$	$y_i$	$x_i$	$y_i$	$x_i$	$y_i$							
		$a_i$		$a_i$		$a_i$		$a_i$		$a_i$		$a_i$		$a_i$								
1	$a_1$	60	0	1	0	0	1	30	30	20	0	10	20	40	40	20	5	5	1	20	20	1
2	$a_2$	30	0	1	90	0	1	30	30	5	60	40	5	10	10	5	6	6	1	1	1	1
3	---	10	10	1	10	10	1	5	5	10	5	0	10	20	20	10	5	5	1	2	2	1
4	---	0	30	1	0	90	1	15	15	10	5	0	10	20	20	10	5	5	1	1	1	1
n	$a_n$	0	60	1	0	0	1	20	20	5	30	50	5	10	10	5	4	4	1	1	1	1
Totals	X Y A	100	100	5	100	100	5	100	100	50	100	100	50	100	100	50	25	25	5	25	25	5

  

OVERLAP INDICES:	
$C_{xy}$ (Eq. 1)	0.10
$C_m$ (Eq. 3)	0.022
$R_o$ (Eq. 7)	0.10
$C_p$ (Eq. 8)	-.74
$L$ (Eq. 11)	0.05
	0.10
	0.89
	3.90
	0.80
	3.90
	1.0
	1.0
	1.0
	1.0
	1.0
	1.0
	1.0
	1.016
	3.256

(1974) elaboration of the competition coefficient (Eq. 13), existing overlap indices ignore such variation. This is not serious, as probably any index could be modified so as to take inequality of  $a_i$  into account. It also is an understandable omission, as in most studies of overlap or association in space, all  $a_i$  are equal (e.g., as with a set of quadrats) while in most studies of dietary overlap it is very difficult to determine  $a_i$ .

It might be argued that an overlap index should only reflect the similarity in resources **used** and not similarity of resource **use**. That is, the index should not reflect the degree to which similarity in resources used influences the likelihood that 2 species will impinge on each other in some way. In that case, one could simply consider all  $a_i$  equal and use Eq. 6 or ignore resource abundance altogether and use Eq. 1. While one may be forced into this position occasionally, it does minimize the biological interpretability of the index values and their suitability as a foundation for discussion of resource utilization strategies, competition, species packing, and so on. Intelligent analysis of these phenomena in nature simply cannot proceed very far in the absence of data on resource state abundances.

The above is not to deny the utility of similarity indices such as Eq. 1, only their appropriateness as indices of niche overlap. If, for example, one was carrying out on a group of species a comparative study of the biochemical or physiological aspects of nutrition, then an index of dietary similarity such as might be provided by Eq. 1 might prove a useful tool.

AN IMPROVED INDEX

The probability or frequency of interspecific encounter, broadly understood, relates well, perhaps exactly, to our intuitive concept of niche overlap. Therefore we propose that niche overlap be defined as **the degree to which frequency of interspecific encounter is higher or lower than it would be if each species utilized each resource state in proportion to its abundance ( $a_i$ )**.

For a given resource state i, the number of interspecific encounters will be proportional to the product of the densities of the 2 species and the size of the resource state. Thus, calculated over all resource states, we have a total number of interspecific encounters proportional to

$$E = \sum_i (x_i/a_i) (y_i/a_i) (a_i) = \sum_i (x_i y_i/a_i); \quad (9)$$

and for the hypothetical situation where the individuals are distributed either randomly or uniformly with respect to the abundances ( $a_i$ ) of the resource states, the number of interspecific encounters would be proportional to

$$E_{uniform} = XY/A. \quad (10)$$

Niche overlap is then calculated as

$$L = E/E_{\text{uniform}} = (A/XY) \sum_i (x_i y_i / a_i). \quad (11)$$

If all resource states are equally abundant (e.g., cases I, II, VI, VII, Table 1), Eq. 11 reduces to Lloyd's (1967) index of "interspecies patchiness":

$$L = n \sum_i (p_{xi} p_{yi}). \quad (12)$$

L assumes a value of zero when no resource state is shared by the 2 species, a value of 1.0 when both species utilize each resource state in proportion to its abundance ( $a_i$ ), and a value  $>1.0$  if each species utilizes certain resource states more intensively than others **and** the utilization functions (or preferences, in a narrow sense of the word) of the 2 species tend to coincide.

The interpretation of L is straightforward. For case I or II the probability of interspecific encounter is only 5% of what it would be if both species were uniformly distributed over the resource states. For case IV the probability of interspecific encounter is 290% (or 3.90 $\times$ ) higher than it would be if both species were uniformly distributed with respect to the abundance of the various resource states (case V).

The difference between cases III and IV (Table 1) can be interpreted as follows: Species X and Y have identical utilization functions in case III, both favoring the less common resource states, so that L exceeds 1.0. In case IV the 2 species are not quite identical in their utilization functions, but the effect of this is more than canceled by the now greater intensity with which they both utilize the less common resource states; and so L exceeds 1.0 by an even greater amount than it does for case III.

The probability of interspecific encounter refers, when the resource states are spatial units of some sort, to the likelihood that 2 organisms will bump into or somehow physically crowd each other during a given time interval. When the resource states are diet categories, the probability of interspecific encounter can be thought of as referring to the likelihood that 2 non-conspecific individuals will attempt to utilize the same food item within some time interval. If the time interval is very short, the interspecific encounter may be envisaged as a physical confrontation. For longer time intervals, the interspecific encounter may consist of one individual arriving 'late' and 'finding' that the food item has disappeared.

A significant property of L is that its value is increased by the presence or inclusion of resource states utilized by neither species. That is, if we were to add 5 resource states ( $i = 6, \dots, 10$ ), with  $x_i = y_i = 0$  for each one, to the data set for case I, the value of L doubles to 0.10. This behavior of L is biologically appropriate and consistent with our definition of niche overlap. It does require, however, consideration of exactly what set of resource states should be used in calculating L and hence exactly what types of comparisons are to be made.

Also, L usually will be influenced by the manner in which a resource axis is partitioned into resource states. As the axis is partitioned more finely, the value of L may increase, decrease or remain unchanged. (In contrast, other indices such as Eqs. 1, 2, 3 and, probably, 7 can only decrease or remain unchanged when the axis is partitioned more finely.) Again, this is consistent with our definition of niche overlap but does necessitate circumspection in procedure and interpretation.

Neither of the above properties poses any special complications to the typical investigation of overlap within a set of species, where both the resource range (i.e., the segment of the resource axis) considered and its partitioning into resource states are held constant and not independently defined for each species pair analyzed.

For good discussions of the pitfalls awaiting the average unwary nichemetrician the reader is referred to Colwell and Futuyma (1971) and Cody (1974, p. 70).

#### INTERSPECIFIC CROWDING

Niche overlap is usually quantitatively nonreciprocal. That is, species X usually impinges on species Y to a different degree than does species Y impinge on species X. Thus we require measures of directional overlap to supplement the information provided by L (Eq. 11). Such measures have been available for some time, the first and most widely used being

$$\alpha_{x(y)} = \sum_i (p_{xi} p_{yi}) / \sum_i p_{xi}^2 \quad (\text{Levins 1968}). \quad (13)$$

This index, the first formally proposed measure of niche overlap, is unacceptable for the same reason Eqs. 2, 3, and 4 were rejected: It depends on the species' distributions over the nonshared resource states. Under certain circumstances and with appropriate modifications (cf. Schoener 1974, Abrams 1976), Eq. 13 may reasonably be used to estimate the competition coefficients for the Lotka-Volterra equations, as Levins (1968) suggested. However, in referring to these coefficients as measures of niche overlap, Levins synonymized 2 distinct concepts and generated a measure of semantic confusion which muddies the literature on these topics even today. Earlier, MacArthur and Levins (1967) used the term "niche overlap" but not in such a way as to make clear the exact meaning they assigned to it.

A very good measure of directional overlap, Lloyd's (1967) 'mean crowding on species 1 by species 2,' appears to have been used little or not at all. It is given as

$$Z_{x(y)} = \sum_i (x_i y_i / X). \quad (14)$$

A more general expression, which allows for variation in resource state size ( $a_i$ ), is Rathke's (1976) 'co-occurrence coefficient,' given as

$$Z_{x(y)} = \sum_i [(x_i y_i)/(X a_i)] \tag{15}$$

= the density of species Y encountered, on the average, by an individual of species X.

The reciprocal measure  $Z_{y(x)}$ , is calculated in analogous fashion. Each could be expressed as a fraction of the value (i.e.,  $Z_{x(y), \text{uniform}}$  and  $Z_{y(x), \text{uniform}}$ ), it would assume if the individuals of both species were uniformly distributed with respect to the sizes ( $a_i$ ) of the resource states. It will be found, however, that

$$\frac{Z_{x(y)}}{Z_{x(y), \text{uniform}}} = \frac{Z_{y(x)}}{Z_{y(x), \text{uniform}}} = L \quad (\text{Lloyd 1967}). \tag{16}$$

I shall refer to  $Z_{x(y)}$  and  $Z_{y(x)}$  as measures of interspecific crowding.

PROBABILITY OF INTRASPECIFIC ENCOUNTER

An appealing feature of Eqs. 12 and 15 as measures of interspecific overlap and crowding is their simple, direct relationship to indices of community diversity based on the probability of interspecific encounter (Hurlbert 1971) and to indices measuring the contagion of clumpedness of single species populations in terms of the probability of intraspecific encounter (Morisita 1959, 1971, Lloyd 1967). Clearly, probability of encounter is a concept which can unify and simplify much of statistical ecology. Here one wishes only to extend the mathematics of patchiness (Lloyd 1967) to the case where  $a_i$  is variable.

Applying the same line of reasoning on which Eq. 10 is based, we can state that the probability or frequency with which individuals of a given species will encounter or crowd each other while utilizing a set of resource states will be proportional to

$$F = \sum_i [x_i(x_i - 1)]/a_i. \tag{17}$$

If the X individuals are imagined to be distributed uniformly over the resource states in proportion to their abundances ( $a_i$ ), then we can calculate

$$F_{\text{uniform}} = X(X - r)/A. \tag{18}$$

The ratio

$$G = \frac{F}{F_{\text{uniform}}} = \frac{A}{X(X - r)} \sum_i \frac{x_i(x_i - 1)}{a_i} \tag{19}$$

then measures the degree to which frequency of intraspecific encounter is higher or lower than it would be if each resource state were utilized in proportion to its abundance ( $a_i$ ). Thus it is a generalized form of Lloyd's (1967) patchiness index. The absolute frequency of encounter (mean density of conspecifics confronting the average individual) is simply the generalized form of Lloyd's (1967) mean crowding, viz.

$$m^* = F/X = \sum_i x_i(x_i - 1)/(X a_i). \tag{20}$$

If the average number of individuals per resource state is small (as is often the case, for example, with

quadrat data) chance or sampling error can bring about large (percentagewise) departures from uniformity of distribution. It will be appropriate in those circumstances to take as our standard of comparison the expected frequency of intraspecific encounter for a randomly (rather than uniformly) distributed population. This is easily done only for those cases where resource state abundance ( $a_i$ ) is constant.

Under the conditions of equal  $a_i$  and random distribution, the expected frequency of intraspecific encounter will be proportional to

$$F_{\text{random}} = n \sum_{k=1}^{\infty} [(e^{-m} m^k)/k!] k (k - 1)/a; \tag{21}$$

where  $m = X/r$ ,

$k$  = number of individuals per resource state, and

$a$  = size or abundance of each resource state.

We can then define

$$G' = \frac{F}{F_{\text{random}}} = \frac{\sum_i [x_i (x_i - 1)]}{n \sum_k \left[ \frac{e^{-m} m^k}{k!} k (k - 1) \right]}. \tag{22}$$

This is exactly equivalent to Lloyd's (1967) patchiness, which he presents in more easily calculable form as

$$G' = \frac{n}{X^2} \sum_i [x_i (x_i - 1)] \tag{23}$$

Equations 19 and 22 yield similar values when X is large relative to r. The error that results from using Eq. 19 as an approximation to Eq. 22 is

$$(G' - G)/G' = -n/(X - n), \tag{24}$$

which indicates that, when all  $a_i$  are equal, X should be at least one hundred times r if we wish this error to be 1% or less. This gives an idea of the error expected when, due to inequality of all  $a_i$  we have no choice but to use Eq. 19.

If one wishes to assume that in some sense an individual can encounter or crowd itself, as when "crowding operates, not by hostile encounters between individuals, but through depletion of some expendable resource" (Lloyd 1967), we obtain as a substitute for Eq. 19

$$G = (A/X^2) \sum_i [(x_i^2/a_i)] , \tag{25}$$

and as a substitute for Eq. 23

$$G' = (n/X^2) \sum_i (x_i^2) , \tag{26}$$

and as substitute for Eq. 20

$$m^* + 1 = \sum_i [x_i^2/(X a_i)], \tag{27}$$

which is the generalized form of Lloyd's (1967) "mean demand." I shall refer to both mean demand and mean

crowding (Eq. 20) as measures of intraspecific crowding.

Equations 22, 23, 25, and 26 are all measures of patchiness and are interpreted in approximately the same terms as was Eq. 19.

Niche breadth is another concept which, mathematically, falls into synonymy with the other concepts discussed here. A widely used measure of niche breadth is

$$B = 1 / \left[ n \sum_i p_{xi}^2 \right] \quad (\text{Levins 1968}), \quad (28)$$

which is simply the reciprocal of patchiness as measured by Eq. 26. In our terminology, it equals (when all  $a_i$  are equal) the frequency of intraspecific encounter expected when all resource states are utilized equally divided by the frequency of intraspecific encounter expected on the basis of the observed distribution. Likewise, a more general measure of niche breadth, one which allows for variation in abundance ( $a_i$ ) of resource state, is the reciprocal of patchiness as measured by Eq. 25, viz.

$$B' = X^2 / \left[ A \sum_i (x_i^2/a_i) \right]. \quad (29)$$

This has the same interpretation as Eq. 28.

Equation 29 can take on values ranging from  $1/n$  (when only a single resource state is used) to 1.0 (when each resource state is utilized in proportion to its abundance). Equation 28 can assume values ranging from  $a_{\min}/A$  to 1.0, where  $a_{\min}$  stands for the abundance of the least abundant resource state.

By scaling Eqs. 27 and 28 it is possible to obtain alternative measures of niche breadth which can take values ranging from 0 to 1 and which may have less cumbersome interpretations. Thus we can define

$$B_a = [B - (1/n)] / [1 - (1/n)] = (nB - 1)/(n - 1), \quad (30)$$

and

$$B'_a = [B' - (a_{\min}/A)] / [1 - (a_{\min}/A)] = (AB' - a_{\min}) / (A - a_{\min}). \quad (31)$$

Both of these may be said to measure the degree to which intraspecific encounters have been minimized as a result of the species utilizing or tending to utilize all available resource states in proportion to their abundances ( $a_i$ ). At one extreme there is the ultimate specialist which utilizes only a single resource state ( $B_s$  or  $B'_s$  equals zero) and ignores the others, and at the other extreme the perfect generalist ( $B_s$  or  $B'_s$  equals unity) which utilizes all resource states without preference.

Note that Eq. 31 yields a value of zero only for a species which specializes on the least abundant resource state; it yields values greater than zero for any species which utilize a single but more abundant resource state. This is simply a specific consequence of taking resource state abundance into consideration;

the unscaled index (Eq. 29) also will yield lower values for specialists on rare resource states than for specialists on more abundant ones.

Schoener (1974) proposed a niche breadth index rather similar to Eq. 29, viz.

$$B_s = X^2 / [A^2 \sum_i (x_i/a_i)^2], \quad (32)$$

which he termed selectivity, even though high values of  $B_s$  indicate low selectivity. The manner in which the variable  $a_i$  is incorporated into  $B_s$  does not permit interpretation of this measure in the same probability of encounter terms we have applied to  $B$  and  $B'$ . On these grounds, Eq. 32 is less satisfactory than is Eq. 29 as the generalized form of Levins' (1968) niche breadth (Eq. 28).

#### SON OF ALPHA

By manipulation of Eq. 12, a formal relationship between niche overlap, competition coefficients, and patchiness can be demonstrated. Specifically,

$$\text{overlap (Eq. 12)} = \left( \begin{matrix} \text{geometric mean} \\ \text{of competition} \\ \text{coefficients (Eq. 34)} \end{matrix} \right) \left( \begin{matrix} \text{geometric mean} \\ \text{of patchiness} \\ \text{indices (Eq. 25)} \end{matrix} \right) \quad (33)$$

$$\frac{A \sum (x_i y_i / a_i)}{XY} = \left[ \frac{\sum (x_i y_i / a_i)}{\sum (x_i^2 / a_i)} \cdot \frac{\sum (x_i y_i / a_i)}{\sum (y_i^2 / a_i)} \right]^{1/2} \cdot \left[ \frac{A \sum (x_i^2 / a_i)}{X^2} \cdot \frac{A \sum (y_i^2 / a_i)}{Y^2} \right]^{1/2}$$

If patchiness is calculated with Eq. 19 or 23, instead of 25, then the above relation is only approximate. The approximation improves as the  $x_i$  increase and as  $r$  decreases.

When data consist of relative abundances, so that  $X = Y = 100\%$ , the index

$$S_{x(y)} = \left[ \sum_i (x_i y_i / a_i) \right] / \left[ \sum_i (x_i^2 / a_i) \right] \quad (34)$$

is simply Levins'  $\alpha_{x(y)}$  (Eq. 13) adjusted for variation in  $a_i$ .  $S_{x(y)}$  would be calculated by letting  $x_i = p_{xi}$  and  $y_i = p_{yi}$ ; and it would measure, for species  $X$  and the hypothetical situation where  $X = Y$ , the probability of interspecific encounter divided by the probability of intraspecific encounter.

When  $S_{x(y)}$  is calculated from data on absolute abundances, it measures, for species  $X$ , the expected ratio of interspecific to intraspecific encounters for whatever the actual values of  $X$  and  $Y$  may be. An alternative expression for Eq. 34 is thus

$$S_{x(y)} = (Y/X) \sum_i (p_{xi} p_{yi} / a_i) / \sum_i (p_{xi}^2 / a_i) = \frac{Z_{x(y)}}{m + 1} \quad (\text{Eq. 27}) \quad (35)$$

Schoener (1974) also has proposed modifications of

TABLE 2. Values of overlap and other indices calculated on the goldenrod (*Solidago* spp.) and flamingo (*Phoenicoparrus* spp.) data presented in the text and in Tables 3 and 4

Data set and species	Niche overlap	Similarity	Competition coefficient	Mean coefficient	Inter-specific crowding	Niche breadth	Mean density	Intra-specific crowding	Patchiness or selectivity
Goldenrods in quadrats									
	(Eq. 12)	(Eq. 1)	(Eq. 13)	(Eq. 2)	(Eq. 14)	(Eq. 28)	X/A	(Eq. 20)	(Eq. 23)
<i>S. juncea</i>	3.06	0.39	0.64	0.36	4.8/m <sup>2</sup>	0.19	1.4/m <sup>2</sup>	6.4/m <sup>2</sup>	4.6
<i>S. nemoralis</i>			0.21		4.3/m <sup>2</sup>	0.07	1.6/m <sup>2</sup>	20/m <sup>2</sup>	12.8
Flamingos on lakes									
	(Eq. 11)	(Eq. 1)	(Eq. 34)	(Eq. 37)	(Eq. 15)	(Eq. 29)	X/A	(Eq. 20)	(Eq. 19)
<i>P. jamesi</i>	1.19	0.35	0.28	0.50	88/km <sup>2</sup>	0.43	135/km <sup>2</sup>	310/km <sup>2</sup>	2.3
<i>P. andinus</i>			0.90		160/km <sup>2</sup>	0.41	74/km <sup>2</sup>	177/km <sup>2</sup>	2.4
Flamingo diets—data set 3a									
	(Eq. 11)	(Eq. 1)	(Eq. 34)	(Eq. 37)		(Eq. 29)			(Eq. 25)
<i>P. jamesi</i>	0.81	0.80	0.49	0.52	...	0.60	...	...	1.66
<i>P. andinus</i>			0.54		...	0.67	...	...	1.48
Flamingo diets—data sets 3a & 3b									
	(Eq. 11)	(Eq. 1)	(Eq. 34)	(Eq. 37)	(Eq. 15)	(Eq. 29)	X/A	(Eq. 27)	(Eq. 19)
<i>P. jamesi</i>	0.81	0.80	0.12	0.52	0.81/kg	0.60	4.0/kg	6.83/kg	1.68
<i>P. andinus</i>			2.18		3.23/kg	0.67	1.0/kg	1.52/kg	1.44

the competition coefficient which, among other things, take into account the variation in  $a_i$ . The simplest form of his modified coefficient is

$$\alpha_m = \left[ \sum_i (p_{xi}p_{yi}/a_i^2) \right] / \left[ \sum_i (p_{xi}/a_i)^2 \right]. \quad (36)$$

When  $X = Y = 100\%$ , which is the only sort of data for which  $\alpha_m$  was intended,  $\alpha_m$  is similar to  $S_{x(y)}$  (Eq. 34). The fundamental distinction between them can be expressed as follows. The  $i^{\text{th}}$  term in the numerator (or denominator) of  $\alpha_m$  is proportional to the number of encounters 'on' a given unit of  $i^{\text{th}}$  resource state; the  $i^{\text{th}}$  term in the numerator (or denominator) of  $S_{x(y)}$  is proportional to the number of encounters 'on' the entire  $i^{\text{th}}$  resource state. Thus, for  $\alpha_m$ , the sum of the numerator terms (or of the denominator terms, is proportional to the probability of encounter averaged over resource states, all weighted equally; for  $S_{x(y)}$  the sum of the numerator terms is proportional to the probability of encounter averaged over all units of resource. Only the latter quantity seems to relate logically and directly to competition and niche overlap. On these grounds, essentially the same ones that caused us to reject  $B_s$  (Eq. 31) as a measure of niche breadth, we conclude that  $\alpha_m$  (Eq. 36) is a less useful competition coefficient than is  $S_{x(y)}$  (Eq. 34):  $\alpha_m$  is not interpretable as the ratio of interspecific to intraspecific encounters.

The geometric mean of the competition coefficients, namely

$$S_m = [S_{x(y)} \cdot S_{y(x)}]^{1/2}, \quad (37)$$

has itself been used as a measure of niche overlap although in a form (Eq. 2) that neglects variation in  $a_i$ . I earlier criticized its use as an overlap index because of its dependence on the species' niche breadths over

nonshared resource states. Note that patchiness indices also are dependent on this characteristic, and that the terms  $\sum_i (x_i^2/a_i)$  and  $\sum_i (y_i^2/a_i)$  formally responsible for these dependences all cancel out when competition coefficients and patchiness are multiplied to obtain niche overlap (Eq. 33).

SOME EXAMPLES

To further illustrate the diverse applications and interrelations of the measures discussed, I will use the 3 sets of data presented below. Various index values calculated on these data are presented in Table 2.

*Example 1. Goldenrods in quadrats.*—Seventy (70) quadrats of 1 m<sup>2</sup> each were selected from a vegetationally heterogeneous 3.5-ha area in Dryden, New York. In each quadrat, the numbers of stems of 2 species of goldenrod, *Solidago juncea* and *Solidago nemoralis* (Compositae), were enumerated. The results were as follows: 45(0,0); 6(1,0); (2,0); 2(3,0); (4,0); (5,0); (6,0); (10,0); (0,1); (1,1); (12,1); (1,3); (7,4); (3,5); (6,5); (5,8); (5,9); (5,11); (4,24); (1,37); the first number inside the parentheses represents the number of *S. juncea* stems, the second the number of *S. nemoralis* stems, and numbers preceding parentheses indicate the number (when >1) of quadrats possessing the particular combination). In this example, the resource states are quadrats, which are all the same size ( $a_i$ ).

These goldenrod data suffer from the same shortcomings that characterize most phytosociological data: they were obtained with artificial sampling units of arbitrarily determined dimensions. The type of analysis applied here (Table 2) will yield biologically more interesting results when the sampling units correspond to natural entities, such as mice on which one might enumerate ectoparasites.

TABLE 3. Abundance of flamingos (*Phoenicoparrus* spp.) on 3 lakes

Lake	Lake area (km <sup>2</sup> )	No. of birds	
		<i>P. andinus</i>	<i>P. jamesi</i>
Laguna Kollpa	1.0	24	343
Laguna Cañapa	0.4	125	100
Laguna Khar Kkota	2.0	102	15

*Example 2. Flamingos on lakes.*—Three small, shallow lakes in the Bolivian altiplano were censused for 2 species of flamingos, *Phoenicoparrus andinus* and *Phoenicoparrus jamesi*, with the results shown in Table 3. These are a subset of censuses of 3 flamingo species on 27 lakes. In this example the resource states are the lakes, which are variable in size ( $a_i$ ).

*Example 3. Flamingo diets.*—The data in Table 4 are artificial but, at least in the case of data set 3a, are of the sort we are presently gathering on these flamingos, which appear to feed solely on diatoms. The resource states are diatom length categories and their abundances in the lake ( $a_i$ ) and in the diets ( $x_i, y_i$ ) are measured as relative or absolute biomass. Data set 3a represents the minimum amount of information sufficient for an appraisal of dietary overlap with respect to 'prey' size. Only if the information in data set 3b is also available can certain other parameters, such as inter- and intraspecific crowding, be estimated.

In this example, it is assumed that there is a lake (or a region) inhabited by 50 *P. andinus* and 200 *P. jamesi*, and that the data on diatom availability and utilization have been obtained by means of a sampling scheme adequate to assure their representativeness.

#### Niche overlap and similarity

The values listed under 'niche overlap' in Table 2 all have essentially the same interpretation. In the example concerning goldenrods, the probability of interspecific encounter is  $3.06\times$  (or 206%) greater than it would be if both species were uniformly or randomly distributed over the 70 quadrats. In the flamingo lake example, the probability of interspecific encounter is

19% greater than it would be if both flamingos were uniformly distributed over the available lake space. In the flamingo diet example, the intensity with which the 2 species impinge on each other's food resources is 19% less than it would be if both species were perfect generalists with respect to diatom size.

The term niche overlap seems appropriate only in the case of the flamingo diets. For the other 2 examples, Eqs. 11 and 12 are better termed measures of spatial overlap, distributional overlap or association.

If one calculates for the goldenrod data a more conventional index of association such as Cole's (1949, Hurlbert 1969), which utilizes only presence-absence data, a strong and positively significant degree of association is demonstrated ( $C_8 = 0.87, P < .01$ ). This result coincides with the high degree of spatial overlap indicated by Eq. 12 (Table 2). However, such coincidence is fortuitous and not to be expected as a matter of course: indices based on abundance data and those based on presence-absence data measure aspects of distribution which are partially independent. Consequently, the two types of indices will often yield results which seem contradictory when considered only superficially (Hurlbert 1969).

Comparison of the niche overlap (Eq. 11 or 12) and similarity (Eq. 1) values calculated for the examples demonstrates the distinctness of these measures. For example, the dietary data show the lowest degree of overlap yet the highest degree of similarity! These contrasts principally reflect the indifference of the similarity index to the existence of unoccupied or unutilized resource states and to variation in resource state abundance ( $a_i$ ). Thus, if we ignore the 45 empty quadrats in example 1 and if we assume the diatom size categories are equally abundant in example 3, we will, on recalculation, obtain overlap values (1.09 and 1.32, respectively) that at least are concordant with the similarity values (unchanged at 0.39 and 0.80 respectively) for these examples.

The mean competition coefficient  $S_m$  (Eq. 37) yields, for the 3 examples, values which are concordant with neither the niche overlap values nor the similarity values. Note that the value of  $S_m$ , unlike the values of

TABLE 4. Size distributions of diatoms available to and consumed by 2 flamingo (*Phoenicoparrus*) species

Diatom length ( $\mu\text{m}$ )	Data set 3a				Data set 3b		
	% of total diatom biomass in lake ( $a_i$ )	% of diet		Diatom biomass in lake (kg)	No. of equivalent birds		
		<i>P. jamesi</i>	<i>P. andinus</i>		<i>P. jamesi</i>	<i>P. andinus</i>	
0-20	10	30	10	5	60	5	
21-40	40	50	30	20	100	20	
41-60	30	20	20	15	40	15	
61-80	10	0	20	5	0	5	
81-100	8	0	10	4	0	4	
$\geq 101$	2	0	10	1	0	1	
Totals	100	100	100	50	200	50	



the individual competition coefficients, is independent of whether it is calculated on relative or absolute abundance data, as is demonstrated by the dietary data (Table 2, example 3).

Nonconcordances such as those demonstrated by these examples contradict the idea that "the particular index used is somewhat arbitrary [i.e., unimportant] since similar qualitative results are obtained with a wide variety of indices" (Pianka 1974). This attitude may also be criticized for the more fundamental reason that it encourages loose formulation of hypotheses.

#### *Interspecific and intraspecific crowding*

The calculation of these properties is appropriate only when data on absolute densities are available. When only relative abundances are known, as in data set 3a (Table 2), then  $A = X = Y = 100\%$ ,  $Z_{x(y)} = Z_{y(x)} = L$ , and  $G$  (Eq. 25) =  $\bar{m} + 1$  (Eq. 27).

Interspecific and intraspecific crowding have very straightforward interpretations and require little further explication. Both are expressed as the mean density of individuals confronting an individual. For example, on the average, a stem of *S. juncea* finds in its quadrat 6.4 other stems of *S. juncea* and 4.8 stems of *S. nemoralis*; on the average an individual of *P. jamesi* will find its food supply being utilized by other individuals of *P. jamesi* and by individuals of *P. andinus* in the ratio of 6.63 to 0.81. Obviously the ratio of interspecific to intraspecific crowding determines the competition coefficient (Eq. 34), exactly or approximately depending on which expression for intraspecific crowding is used.

The ratio of the interspecific crowding values for two species is the reciprocal of the ratio of their abundances. That is,

$$Z_{x(y)}/Z_{y(x)} = Y/X; \quad (38)$$

a less exciting result is difficult to imagine.

#### *Niche breadth and patchiness*

As niche breadth and patchiness are reciprocals of each other, generally 1 concept (and measure) will suffice in any given situation. For the examples (1 and 2) which treat of spatial distributions, patchiness seems the more relevant one. However, if the term niche breadth is replaced by one such as *distributional uniformity*, Eqs. 28 and 29 (or possibly Eq. 30 or 31) take on a new appeal. It is reasonable to select whichever concept and index possesses mathematical properties and an interpretation most suitable to the study at hand. For certain kinds of statistical or theoretical treatment, it may be convenient to employ an index that can assume values only between zero and unity.

Patchiness calculated on dietary data (example 3) is a measure of the degree of selectivity or specialization and perhaps would be better termed as such. It is still

the reciprocal of niche breadth and superfluous if one prefers to deal specifically with the latter. However the patchiness or selectivity indices have a very simple, direct interpretation and may prove the more useful approach in many situations. For example, the selectivity value of 1.66 for *P. jamesi* indicates that the probability of intraspecific encounter or the probability of 2 *P. jamesi* individuals attempting to utilize the same unit of food simultaneously is 66% higher than it would be if *P. jamesi* were the perfect generalist.

Niche breadth values for flamingo diets yielded by Eq. 29, and naturally the selectivity values also, are rather similar. The 2 flamingo species are located at about the same spot along the specialist-to-generalist spectrum. The perfect generalist would utilize each resource state in proportion to its availability, i.e., would exhibit no preferences, and would be assigned a niche breadth of 1.0 by Eq. 29.

If the variable abundances of the resource states are ignored and niche width is calculated with Eq. 28, we obtain values of 0.44 for *P. jamesi* and 0.83 for *P. andinus*. These might be taken as evidence that *P. jamesi* is much more of a specialist than is *P. andinus*. However, Eq. 28 gives great weight to the fact that *P. jamesi* does not utilize diatoms  $>60 \mu\text{m}$  in length. Equation 29, on the other hand balances this nonutilization against the fact that diatoms  $>60 \mu\text{m}$  are relatively uncommon (only 20% of total) anyhow; hence, their nonutilization is not taken as strong indication of a contracted niche or high degree of specialization.

For the dietary data, selectivity has been calculated both by Eq. 19 and by Eq. 25 in order to demonstrate the slight difference between the values they yield (Table 2). Equation 25 could have been used for both analyses of these data, and may be more appropriate for dietary data in general. Equation 19, however, seems more appropriate for data on spatial distributions.

#### DISCUSSION

A limited repertoire of mathematical expressions suffices for the quantification of many spatial and resource utilization characteristics of 1- and 2-species populations. Most useful may be: Eq. 11, niche overlap; Eq. 15, interspecific crowding; Eqs. 19, 23, and 25, patchiness; Eqs. 20 and 27, intraspecific crowding; and Eqs. 28, 30, and 31, niche breadth. All are expressed in terms of the frequency of interspecific or intraspecific encounters.

While these measures are simple to calculate, their use is less straightforward. In most cases, the calculated value of an index and its interpretation will depend on: how resource states are defined; whether they are arbitrary units or discrete natural entities; whether or not "empty" resource states are excluded from the analysis; and whether a uniform or a random distribution is taken as a standard of comparison. These in turn depend on whether one is concerned

with spatial patterns or resource utilization (not that these are mutually exclusive phenomena) and on the kinds of comparisons that are to be made. The wide variety of circumstances in which these indices can be applied makes it impossible to present any concise set of guidelines for their use. *Caveat calculator* is the only universal one. Others are discussed by Lloyd (1967) and Colwell and Futuyma (1971).

Neither the older indices of overlap nor those developed here measure the intensity of competition, except perhaps under special circumstances. The lack of correspondence between overlap and competition was clearly discussed by Colwell and Futuyma (1971) and the reasons for it are more or less self-evident. The noncorrespondence represents a severe restriction on certain theoretical and observational (nonexperimental) approaches to the study of competition: and in desperation, some workers partial to these approaches simply have ignored the noncorrespondence while others have invoked whatever assumptions were required to convert noncorrespondence into correspondence.

In view of the above, it is appropriate to reiterate 3 principal reasons why overlap indices fail as measures of competition. First, the resource considered may not be so scarce as to be limiting, in which case even complete overlap will not result in competition. Second, intensity of competition is not likely to be a smoothly increasing function of densities, e.g., as in directional overlap (Eq. 15), even when resources are limiting. And third, the competitive interactions most influential in determining overlap of realized niches will be those that have taken place prior to the moment of observation. The stronger these prior interactions are, the smaller the degree of overlap that will be observed, other things being equal. Of course, if we determine the resource utilization spectrum of each species in the absence of the other and calculate overlap of their fundamental niches (with respect to a given axis), this third argument is invalidated. However, most studies to date have calculated only overlap of realized niches.

The indices presented have been defined for sampling universes or complete collections. Strictly proper procedure requires that when any of these indices is calculated for a sample, this be done using unbiased sample estimators. Such an approach is severely constrained by the frequent difficulty or impossibility of deriving such estimators and their standard errors; the latter might be desired for calculation of confidence intervals or tests of significance. However, given the nature of the data, such statistical proprieties often will be unwarranted; in a sense, we are saved by our imprecision. For example, individuals rarely are selected at random, neither in studies of spatial patterns nor in studies of, e.g., dietary overlap, even though the sampling units containing them may be. Also, the delimitation of the universe from which the samples

are drawn is usually subjective. In this context, it is overly fastidious to require adherence to conventional good form. Where it is desirable to calculate confidence intervals or to perform significance tests, one can do no better than to follow the advice of Horn (1966): obtain replicate sets of samples, calculate the desired index (in parametric form) for each set, and calculate the necessary variances, confidence intervals, or test statistics from these replicate values.

Finally, it may seem that indices that attempt to express characteristics of spatial distributions in terms of crowding or encounters are based, to a greater extent than are more abstract indices, on restrictive assumptions relating to behavior and the other factors that, in nature, will actually determine encounters. This is not the case. The indices proposed here are based only on the same implicit assumption underlying essentially all statistical approaches to these and related problems: the assumption that all individuals in a given species are equivalent. In defining  $L$ ,  $Z_{x(y)}$ , etc., we simply make explicit a corollary of this assumption: that the individuals are equivalent in their behavior.

#### ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation, Grant No. DEB76-02888. The author is indebted to B. D. Collier, R. K. Colwell, and G. W. Cox for their criticisms and suggestions.

#### LITERATURE CITED

- Abrams, P. 1976. Limiting similarity and the form of the competition coefficient. *Theor. Popul. Biol.* **8**:356-375.
- Cody, M. L. 1974. Competition and the structure of bird communities. Princeton Univ. Press, Princeton, New Jersey. 318 p.
- Cole, L. C. 1949. The measurement of interspecific association. *Ecology* **30**:411-424.
- Colwell, R. K., and D. J. Futuyma. 1971. On the measurement of niche breadth and overlap. *Ecology* **52**:567-576.
- Goodall, D. W. 1973. Sample similarity and species correlation, p. 106-156. In R. H. Whittaker [ed.] *Ordination and classification of communities*, Handbook of vegetation science, Part V, R. Tuxen [ed.-in-chief]. Junk, The Hague.
- Horn, H. S. 1966. Measurement of 'overlap' in comparative ecological studies. *Am. Nat.* **100**:419-424.
- Hurlbert, S. H. 1969. A coefficient of interspecific association. *Ecology* **50**:1-9.
- . 1971. The non-concept of species diversity: A critique and alternative parameters. *Ecology* **52**:577-586.
- Levins, R. 1968. Evolution in changing environments: Some theoretical explorations. Princeton Univ. Press, Princeton, New Jersey. 120 p.
- Lloyd, M. 1967. 'Mean crowding.' *J. Anim. Ecol.* **36**:1-30.
- MacArthur, R. H. 1972. Geographical ecology. Harper and Row, New York. 269 p.
- , and R. Levins. 1967. The limiting similarity, convergence, and divergence of coexisting species. *Am. Nat.* **101**:377-385.
- May, R. M., and R. H. MacArthur. 1972. Niche overlap as a function of environmental variability. *Proc. Natl. Acad. Sci.* **69**:1109-1113.
- Morisita, M. 1959. Measuring of interspecific association and similarity between communities. *Mem. Fac. Sci. Kyushu Univ. Ser. E* **3**:64-80.

- . 1971. Composition of the  $I_g$ -index. *Res. Popul. Ecol.* **13**:1-27.
- Pianka, E. R. 1973. The structure of lizard communities. *Annu. Rev. Ecol. Syst.* **4**:53-74.
- . 1974. Niche overlap and diffuse competition. *Proc. Natl. Acad. Sci. USA* **71**:2141-2145.
- Rathke, B. J. 1976. Competition and coexistence within a guild of herbivorous insects. *Ecology* **57**:76-87.
- Renkonen, O. 1938. Statistisch-ökologische Untersuchungen über die terrestrische Käferwelt der finnischen Bruchmoore. *Ann. Zool. Soc. Zool. Bot. Fenn. Vanamo* **6**(1):1-231.
- Schoener, T. W. 1970. Non-synchronous spatial overlap of lizards in patchy habitats. *Ecology* **51**:408-418.
- . 1974. Some methods for calculating competition coefficients from resource utilization spectra. *Am. Nat.* **108**:332-340.
- Whittaker, R. H., and C. W. Fairbanks. 1958. A study of plankton copepod communities in the Columbia Basin, southeastern Washington. *Ecology* **39**:46-65.