A comparison of association indices

SARA J. CAIRNS* & STEVEN J. SCHWAGER‡
Field of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853, U.S.A.
†*Biometrics Unit, Cornell University, Ithaca, New York 14853, U.S.A.*

Abstract. Indices used to measure the frequency of association between individuals in fission/fusion societies are frequently borrowed from ecological studies of species association without adequate justification. In this paper several such indices, under specific conditions likely to be encountered in field studies of animal behaviour, are examined. Each of three indices commonly found in the literature is shown through simulations to be accurate over only a narrow range of possible sampling biases. As an alternative approach, examples are given of the derivation of a maximum-likelihood estimator based on two simple models. Both models assume that a constant proportion of existing subgroups is located. In the first model all combinations of subgroups are equally likely to be located, while in the second model the probability of locating one subgroup is allowed to differ from that of the others. The maximum-likelihood estimators are shown to be less biased and to have lower variance than the other three indices under the assumptions of the models.

The frequency with which two individuals associate is basic to all other aspects of their social interactions. There are, however, serious problems with the methods currently used to measure individual association. Association is defined here to mean the frequency with which two individuals are present in the same social group at the same time. In particular, this paper will be concerned with species whose social organization is of the fission/fusion type. An example of such a species is the lion, *Panthera leo*: a pride of lions is a stable, closed group of related individuals, but at any moment the pride occurs in scattered groups whose membership changes from day to day (Schaller 1972). For any species, when group membership changes at frequent intervals, the level of association between individuals may vary widely. For instance, individuals may change groups at random, showing no preference for particular companions. Alternatively, pairs or small cliques may have such stable bonds that they form indivisible subunits of the larger, changeable groups. More likely than either extreme is a probabilistic relationship, where a pair that tends to associate may on any given day be either together or in separate groups.

A wide variety of species occur in groups with changeable group membership. These include ungulates, carnivores, primates, bats and birds. For many species it is still unclear just what frequency of association is typical of different age–sex classes. The only descriptions of association available are often based on subjective impressions acquired during research focused on other aspects of animal behaviour. It is highly desirable, however, to measure association quantitatively. This is the only way to avoid observer bias, to give equal weight to the more and less conspicuous members of the population, and to make useful comparisons between populations or species. It is also particularly important to quantify variables (such as association) that can be measured only by observations occurring over long periods of time. Subjective judgements are likely to be unduly influenced by recent events or by periods of time during which (even unrelated) events occurred that made a strong impression on the observer.

Many recent attempts have been made to quantify frequency of association in a variety of species. The most common technique used is to search an area, noting the identities of individuals in all groups encountered. Another method is to check specific locations where groups are known to form, and another is to attract social groups with food baits. Examples include walking along a beach and recording the presence of banded individuals in different flocks of sanderlings (*Calidris alba*, Myers 1983), checking group composition in specific trees used as day roosts by vampire bats (*Desmodus rotundus*, Wilkinson 1985), and observing arrivals of chickadees, *Parus atricapillus*, at bird feeders

(Ficken et al. 1981). The data collected are used to score pairs of animals as together (in the same group) or separate (each animal in a different group) at each sample. In captive studies it may be possible to take instantaneous samples (Altmann 1974) of all subjects. In field studies one sample is likely to consist of sightings of several groups during a period of several hours. Different groups will not be located at the same instant, but changes in group membership during the time it takes to locate them are unlikely. It will be difficult or impossible to locate all existing groups at each sample for non-captive populations.

Many different indices of association can be calculated from the same basic set of data. A formula used to estimate association should have the following properties. (1) It should use data that it is feasible to collect. (2) The expected value of the index should be equal or close to the true frequency of association between pairs. (3) The index should have low variance, allowing different levels of association to be distinguished even with small sample sizes. (4) It should estimate a variable whose absolute level can be compared between populations. (5) It should be relevant to the subjects; that is, be based on behaviours that vary according to the degree of association between individuals.

The question of relevance to the subject is basic both to the initial decision of how to define association (e.g. at what distance animals are considered to be in separate groups) as well as to the interpretation of the final index of association (see Sailer & Gaulin 1984 for a discussion of whether an animal's relative or absolute level of co-occurrence with other individuals should be used in evaluating its social network). These topics will not be discussed here: our purpose in this paper is to evaluate the bias, precision and comparability of different association indices.

In situations where the researcher does locate all individuals at each sample (typically only in studies of captive populations), most of the above properties are satisfied by calculating the number of samples in which a pair is together divided by the total number of samples. Whenever it is impossible to locate all individuals in each sample, however, the possibility arises that the samples will be biased. The probability of locating an animal may be related to the number or type of its companions, its reproductive condition, or many other factors that in turn are related to its pattern of association with other members of the population. In such cases,

special attention must be paid to the suitability of the index chosen to estimate association frequency.

In the next section we describe the indices most frequently used in behavioural studies, make explicit their underlying assumptions, and point out possible sources of error and confusion. A method of 'custom-designing' association indices will then be described, and finally recommendations will be offered for future studies of individual association.

Table 1. Definitions of notation for types of samples and total sightings of individuals

Definitions

Types of samples

$N[E]$ denotes the number of samples in which event E occurs

$x = N[\text{A and B are located together in one group}]$

$y = N[\text{A and/or B are located in separate groups}]$

$\quad y = y_a + y_b + y_{ab}$

$\quad y_a = N[\text{only A is located}]$

$\quad y_b = N[\text{only B is located}]$

$\quad y_{ab} = N[\text{both A and B are located separately}]$

$z = N[\text{neither A nor B is located}]$

$n = x + y + z = \text{total number of samples}$

Total sightings

Of A: $n_a = x + y_a + y_{ab}$

Of B: $n_b = x + y_b + y_{ab}$

Of A without B: $T_a = y_a + y_{ab}$

Of B without A: $T_b = y_b + y_{ab}$

Of A and B together: $T_t = x$

Estimated probability

Of locating A: $\hat{P}(A) = n_a/n$

Of locating B: $\hat{P}(B) = n_b/n$

Of locating both together: $\hat{P}(AB) = x/n$

CURRENT INDICES OF INDIVIDUAL ASSOCIATION

Several indices are now used in studies of pairwise association. All have as a numerator the total number of samples when two specific animals A and B are located together. The denominators differ: one uses the total sightings of A and B; one sums all sightings of A and B together, A without B, and B without A; one scores samples (versus sightings) as together versus separate; and one is based on the probability of sighting A or B.

The differences between them are only clear when all are translated into a common notation.

Table II. Formulae of association indices

Index	Traditional formula	Translation to xyz notation
(a) Half-Weight	$\dfrac{x}{\frac{1}{2}(n_a + n_b)}$	$\dfrac{x}{x + y_{ab} + \frac{1}{2}(y_a + y_b)}$
(b) Twice-Weight	$\dfrac{T_t}{T_t + T_a + T_b}$	$\dfrac{x}{x + 2y_{ab} + y_a + y_b}$
(c) Simple Ratio	$\dfrac{x}{x + y}$	$\dfrac{x}{x + y_{ab} + y_a + y_b}$
(d) Square Root	$\dfrac{\hat{P}(AB)}{\sqrt{\hat{P}(A)\hat{P}(B)}}$	$\dfrac{x}{\sqrt{(x + y_a + y_{ab})(x + y_b + y_{ab})}}$

Notation as in Table I.

The notation to be used in this paper is given in Table I, and the formulae for the indices described above are given in Table II. The indices will be referred to by names intended to aid in remembering which formula is involved: indices (a)–(d) in Table II will be called the Half-Weight Index, the Twice-Weight Index, the Simple Ratio Index, and the Square Root Index, respectively.

The Half-Weight Index, which has also been called Dice's, Sorensen's and the Coherence Index, is by far the most popular with behaviourists (e.g. Knight 1970; Schaller 1972; Houston 1974; Morgan et al. 1976; Leuthold 1979; Ficken et al. 1981; Underwood 1981; Clutton-Brock et al. 1982; Hayaki 1983; Penzhorn 1984; Wilkinson 1985). The Twice-Weight (or Coincidence) Index has been used by Ekman (1979), Guinness et al. (1979), Myers (1983) and Sailer & Gaulin (1984). The Square Root Index (Probability Coefficient) was derived by Lott & Minta (1983) in order to measure association frequency in American bison, *Bison bison*. The Simple Ratio is usually used in captive studies, although Guinness et al. (1979) and Clutton-Brock et al. (1982) have used it in field studies.

When both members of a pair are located in every sample ($z = y_a = y_b = 0$) all indices except the Twice-Weight are identical, with the denominator equalling the total samples taken. On the other hand, if $y_{ab} = 0$ (A and B are never both found when they are in separate groups), then the Twice-Weight and Simple Ratio indices are identical. This would occur if the sampling process were stopped as soon as one member of the pair was located. For a given pair this could be a reasonable way to

sample, since after finding A in a group that does not contain B, no further information is gained by actually finding B. Wherever B is, B is not with A. Usually, however, association for many pairs is being measured, and as many groups as possible are located. The vast majority of field studies include both kinds of sightings when A and B are separate: only one of the pair is located (y_a or y_b), or both are (y_{ab}). Each of the four indices given in Table II will give a different estimate of level of association in such cases. The choice of which index to use basically involves deciding what weight should be given to the two types of separate sighting. There is no indication that this fact has been recognized by researchers measuring association.

The Half-Weight and Twice-Weight indices are monotonic functions of one another, and thus will always lead to the same ranking of association over a set of pairs. That is, if the Half-Weight Index for pair 1 exceeds the Half-Weight Index for pair 2, then the Twice-Weight Index for pair 1 will exceed the Twice-Weight Index for pair 2, and conversely. The Simple Ratio Index, however, is not a monotonic function of either the Half-Weight or the Twice-Weight Index. The conclusions of any statistical test based on an index of association may therefore depend on which index was applied to the data. Comparisons of association indices between studies also require careful choices, since the wrong index will either over- or underestimate the true level of association. Whether an index is accurate depends on the sampling procedure. It is possible that two studies using the same index should not be
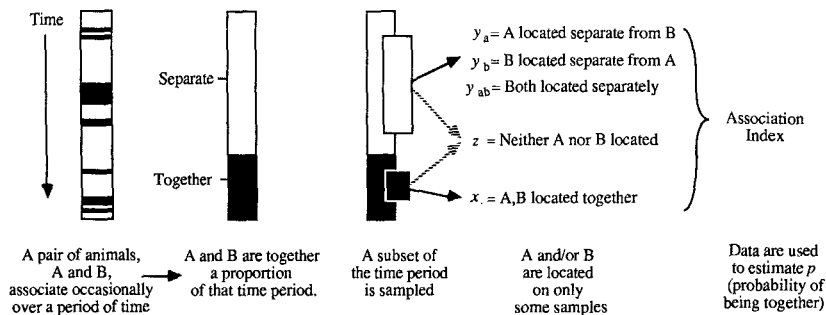
Figure 1. Estimating association frequency for pairs of animals in fission/fusion societies.

compared, since the index might be appropriate in one case but not in the other, due to different sampling procedures. On the other hand, two studies using different indices could be compared if the indices were estimating the same parameter and each was unbiased under the particular circumstances of that study.

Only one of the four indices in Table II, the Square Root Index, was developed specifically for the purpose of measuring individual association. Its derivation by Lott & Minta (1983) involves a potential error that illuminates the difference between scoring individual association versus species association (the purpose for which the Half-Weight and Twice-Weight indices were originally derived). Lott & Minta (1983) used a 2×2 contingency table to develop formulae to estimate probabilities of different types of sightings of a pair of individuals, A and B. The first step of their derivation was the statement that 'If a pair of individuals are randomly associating, then they are occurring in groups independently of each other. In that case the following relation should hold $P(AB) = P(A)P(B)$' (where $P(A)$, for example, is the probability that a sighting of individual A will occur). The contingency table's four cells represent counts of discrete events: both A and B may co-occur $(A + B +)$, A may occur without B $(A + B -)$, B may occur without A $(A - B +)$ or neither may occur in the sample $(A - B -)$. The estimated probability of A occurring with B, $P(AB)$, would be equal to the number of counts in the $A + B +$ cell divided by the total samples. The estimated probability of A occurring in any sample, $P(A)$, would be the sum of counts for the $A + B +$ and $A + B -$ cells, divided by the sample size. However, it is possible in multigroup samples to locate both A

without B and B without A in a single sample (e.g. in two separate groups encountered during a search of the study area). This means that one sample could appear in two cells of the contingency table $(A + B -$ and $A - B +)$ and the sum of all four cell counts would then be greater than the total number of samples. To take the extreme case, suppose that A and B are located on every sample (sometimes together, sometimes in separate groups). Then $P(A) = P(B) = 1 \cdot 0$, and $P(A)P(B) = 1 \cdot 0$. Clearly it is erroneous to conclude that $P(AB)$, the probability of their occurring together, is then equal to $1 \cdot 0$.

The problem created by multigroup samples could be avoided by scoring only the first sighting of a given pair during each sample period. In practice, this would be somewhat complicated. Suppose, for example, that only two groups are located, one consisting of A alone, the other containing B and C. For the pair A–B, this sample could be scored as a sighting of 'A without B', but the sighting of B in the second group would then have to be ignored (not considered to be a sighting of 'B without A') to avoid double counting in the contingency table. Total sightings of individuals would have to be recorded separately from data on pairwise associations. None of the papers cited here clearly use such a scoring technique, since statements such as 'only one observation was recorded per day and pair of individuals' (Ekman 1979) are ambiguous. This could mean that the procedure described above was followed, or that if A was located separately from B but was then found to have joined B later during the same sampling period, that only the first, separate, sighting would be scored.

This situation differs from the case where A and B are species, and a sample consists of a single area

in space. Species A can be said to occur independently of species B if the knowledge that A occurs in a given plot of land provides no information about the probability that B occurs on that same plot of land (in which case $P(A)P(B) = P(AB)$). It is impossible for A and B both to be located on that plot without being considered to co-occur (y_{ab} must equal 0: if both species are found, the sample belongs in cell $A + B +$ of the contingency table). The basic difference between the measurement of species and individual association is the number of groups (or plots) in one sample. If two plots of land are investigated, species A may be located in both plots: the second plot is a separate sample of species association. If, however, at a given moment a population is split into two groups, locating the second group is not equivalent to taking a separate sample of individual association. If individual A was in the first group located during a search of an area, he or she will not be in any other group that exists at that time. Locating many groups simply increases the information available on association patterns for many different pairs during one sample.

There are two important implications of the potential for scoring both A and B separately when they are individuals rather than species. The first is that indices suitable for species association, where the y_{ab} case is an impossibility, may be highly inaccurate in measuring individual association if used with data where more than one group was located per sample period. The second is that the validity of the method used cannot be assessed unless this aspect of scoring is reported explicitly. Both of these warnings apply to the Square Root Index.

Since the scoring of species association is not directly comparable to scoring individual association, ecological indices should not be borrowed without explicit justification. Yet typically the only explanation given for the use of the popular Half-Weight Index is a citation of one of two classical papers that proposed its use for the measurement of the association of species (Dice 1945; Sorensen 1948). The next section will clarify the conditions under which the Half-Weight, Twice-Weight and Simple Ratio indices will accurately estimate association level.

Comparison of Indices

Figure 1 summarizes the process of estimating association frequency between a pair of individuals. The following are points at which the researcher can influence the accuracy of the results.

(1) Deciding when and how often a sample will be taken. The timing of the sample must be determined independently of the subjects' behaviour, and samples should be spaced widely enough to minimize autocorrelation between samples. Larger numbers of samples will increase the power and accuracy of estimates of association, but if made too close together the observations will not be independent: a second sample will not add much information to the first, since it could have been predicted from the first.

(2) Choosing a method of sampling. Ideally, samples would be taken completely at random with respect to whether two animals were together or separate. There are many ways in which a sampling method might introduce a bias, however. For example, if driving through bushy terrain, it may be easier to locate large groups than small ones. If two animals tend to associate only in large groups, they will then be more likely to be found when together than when separate. Biases such as this may be difficult or impossible to detect. When neither one of a pair is located on a given sample, there is no way of knowing whether they are together or separate, thus there is no way to determine whether the failure to locate them was due to a factor that is correlated with the probability of their being together. It may be necessary to use two sampling methods, so that a bias introduced by one can be detected by the other.

(3) Choosing an index that will summarize the data. If the sample is also the population of interest (no inferences are to be made about times when the animals were not observed), then the Simple Ratio Index gives an exact answer to the question of how often a pair associated. Usually, however, the goal will be to estimate the frequency of association over a time period that was only occasionally sampled, using samples that include occasions when neither one of the pair was located. This is the situation where choice of an index becomes critical.

The Half-Weight and Twice-Weight indices in Table II weight separate sightings (y_a, y_b and y_{ab}) by either doubling or halving different variables. Such weights are usually used in equations to correct explicitly for known or assumed bias in the sample. The effect of the weights in these indices, however, will depend on the relative magnitude of x and y, as well as that of y_{ab} to y_a and y_b. These in turn will be

determined by the actual proportion of time spent together, and by any biases in the probability of locating individuals introduced by the sampling method.

The effect of these two factors was explored using a computer simulation of association. The simulation mimicked a field sample of association between two individuals, using random numbers to determine what was actually 'observed' at each 'sample' (a 'sample' actually being one execution of a set of calculations). The pair was considered to have been together during the proportion p of the total time periods sampled. Different probabilities of observing the pair were used when they were together versus when they were separate: the relative values of these probabilities ranged from 2:1 in favour of finding at least one member of the pair when they were separate to 2:1 in favour of finding the pair when they were together.

For each time period in which the pair was together, a random number was calculated and compared to the probability of locating the pair when together. If this probability was greater than the random number, x was incremented by 1 (the pair was observed during the sample), otherwise z was incremented. The same operations were performed for each time period when the pair was separate, determining whether at least one of the pair was located. If so, another random number determined whether both were seen or only one, thus whether y_{ab} was incremented or y_a or y_b.

At the end of each simulation run, or repetition, values for x, y_a, y_b, y_{ab} and z had been generated that summed to the total number of samples (fixed at 60 for all runs). These were used to calculate three indices: the Half-Weight, the Twice-Weight and the Simple Ratio. An average value for each index was calculated based on 1000 repetitions of the simulation, holding all variables constant except for the random numbers. (The simulation was written in PL/I and run on an IBM 370 using the IBM Uniform Random Number Generator with an odd 9-digit seed.)

The proportion of time periods spent together by A and B (p) was either 0·25 (15 out of 60), 0·50 (30 periods), or 0·75 (45 periods). Relative probabilities of locating at least one member of the pair when together versus separate were either 0·25 versus 0·50 (1:2, or twice as likely to locate when separate as when together), 0·50 versus 0·75 (2:3), 0·50 versus 0·50 (1:1, or equally likely), 0·75 versus 0·50 (3:2), or 0·50 versus 0·25 (2:1, or twice as likely

when together as when separate). On samples when the pair was separate and at least one member was located, the probabilities of locating only one versus both were either 0·33 versus 0·67 (1:2), 0·50 versus 0·50 (1:1), or 0·67 versus 0·33 (2:1).

The effect of the third variable, the relative probability of locating only one individual versus both when they were separate, was found to be negligible compared to the other two. The results given here are those for which the y_{ab}-type and y_a- or y_b-type sightings were equally likely (0·50 versus 0·50).

The average deviation of each index from the true proportion of time periods A and B spent together is presented in Fig. 2 (points above zero are overestimates, points below zero are underestimates). The Half-Weight is least biased if pairs are more likely to be scored when separate than when together. The Twice-Weight is least biased when the sampling is biased in favour of sightings of A and B together, while the Simple Ratio Index is least biased when the sample is random.

Each of these three indices is thus best suited for a different sampling bias. Any one index is less biased than the other two for only a narrow range of bias, while none of the indices can be expected to perform well if the sampling bias is greater than 2:1 in either direction. Large errors in estimation can occur even with relatively small sampling bias: the Twice-Weight Index, for example, with a bias of 3:2 in favour of separate sightings, estimated on average that a pair spent 31% of the samples together when they were actually together on 50%. One conclusion that can be drawn is that it is necessary, before choosing one of these indices, to assess what biases might be introduced by the sampling method. As mentioned earlier, it may be necessary to use two different methods concurrently and compare them to each other.

The results shown in Fig. 2 indicate which of the three indices would be better than the others for general directions of bias, but do not provide any guarantee of the absolute accuracy for a specific situation. An alternative approach to choosing an index simply because it is empirically better than another is to develop an index specifically designed to deal with a given sampling method. This can be done by making explicit assumptions about the sampling procedure in order to develop a mathematical model, which can then be used to derive an estimate of association level with known statistical properties. Appropriate assumptions would vary
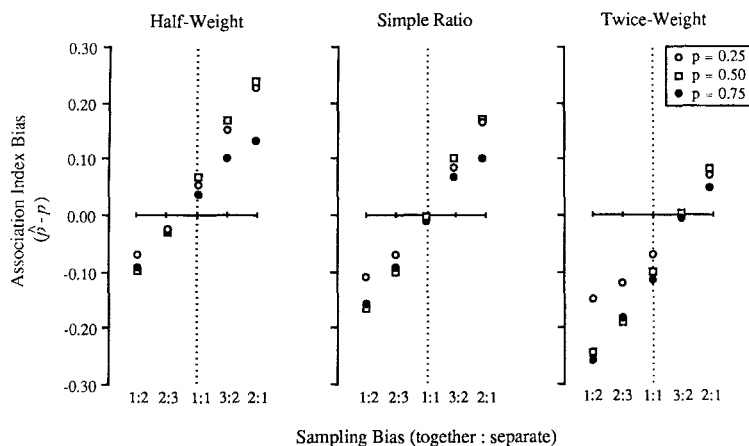
Figure 2. Bias of three association indices as a function of association frequency (p) and sampling bias. Dotted line indicates no bias (equally likely to locate at least one of the pair when they are together as opposed to when the pair is in two separate groups). Actual probabilities of locating the pair when together versus separate are given in the text. Each point is the average value of 1000 repetitions of a simulation.

from case to case, but even very general ones can produce useful estimates. In the next section two models are described, both to present specific indices and to demonstrate the technique of developing a maximum-likelihood estimator.

MAXIMUM-LIKELIHOOD ESTIMATES OF ASSOCIATION

The models to be described here simply consist of a precise set of assumptions about how a collection of observations might be generated. These assumptions can be used to develop formulae for the expected values of each type of observation, and these formulae can then be used to derive the likelihood function for a specific statistic (which may be any function intended to estimate the true value of a population parameter). The likelihood function is an equation that describes how the probability of observing a specific value of a statistic (e.g. the total number of heads actually obtained in 10 coin tosses) varies as a function of the true value of the parameter generating the observations (e.g. the probability of a head on any one toss). If the likelihood function meets certain

mathematical conditions, the parameter value that maximizes the likelihood function can be found routinely. This maximizing value is the maximum-likelihood estimator of the true parameter value. (In the case of coin tosses, the maximum-likelihood estimator for the probability of a head on any one toss is h/n, where h is the number of heads observed in n tosses.) Maximum-likelihood estimators are intuitively easy to understand, and are usually possible to derive. Additionally, in the case of asymptotically large sample sizes under very general conditions, they can be shown to be at least as good as any other estimator (Mood et al. 1974, page 359).

Two models will now be proposed, whose maximum-likelihood estimators will be derived and then compared to the indices found in the literature. These models are developed not only to propose possible indices, but also to explore how the indices now being used behave under specific sampling conditions. The maximum-likelihood estimator derived for a particular model will be shown to be a better estimator for that model (having less bias and lower variance) than any of the currently used indices of association. The maximum-likelihood estimator for both models, however, uses in its computation a parameter (the

number of existing social groups) that can only be estimated in an acutal field situation. Depending on how accurately this could be estimated, the maximum-likelihood estimator may or may not be a practical choice of index.

Model I: Locate *j* out of *k* Groups

The basic assumption of the first model to be considered is that at the time of each sample the population is divided into *k* groups, *j* of which are located by the observer. The values of *j* and *k* are assumed to be constant for this simple model and each possible set of *j* groups is equally likely to be obtained. The goal is to estimate *p*, the probability that a given pair, A and B, will be together in the same group at any arbitrary point in time. One of the advantages of this model is that it does not require specifying the size of the *k* groups.

From these assumptions it is possible to develop formulae for the probability of each type of sighting (x, y_a, y_b, y_{ab} and z). For example, there are $\binom{k}{j}$ ways of choosing *j* out of *k* objects when sampling without replacement (where $\binom{k}{j} = k!/(k-j)!j!$); this is the total number of possible outcomes for one sample. Locating both A and B when they are separate means locating the two groups containing them as well as $j-2$ other groups out of the $k-2$ that do not include A or B. Thus the probability of a y_{ab}-type sighting is

$$\binom{k-2}{j-2} \Big/ \binom{k}{j} = j(j-1)/k(k-1)$$

A formula for the probability of each of the types of sighting can be determined in a similar fashion. These formulae determine the likelihood function for *p* given the observed values of x, y_a, y_b, y_{ab} and z. The likelihood function can then be used to derive the maximum-likelihood estimator. (Details are given in Appendix I.) The resulting formula for $\hat{p}$, the maximum-likelihood estimator of *p*, is

$$\hat{p} = [w + \sqrt{w^2 + 4nj(k-j-1)x}]/2nj$$

where
$$w = (2j-k+1)x + (j-k+1)y + jz$$

Note that when $z=0$ (as in a captive study, where all pairs are located at each sample), this formula simplifies to the Simple Ratio Index ($x/(x+y)$). Also, if it were possible to know that all but one group had been located ($j=k-1$), it would simplify to $(x+z)/(x+y+z)$.

The maximum-likelihood estimator can be compared to other indices using a simple numerical example. Suppose that two individuals A and B spend 50% of their time together ($p=0.50$), $n=20$ samples are taken and $j=3$ out of $k=5$ groups are located at each sample. Using the formulae given in Appendix I, the expected values for types of sightings are $x=6$, $y=9$ and $z=5$. The different types of separate sightings (y_a, y_b and y_{ab}) each have a value of three. Calculating the indices using these values, the Twice-Weight Index (0·33) and Simple Ratio Index (0·40) underestimate *p*, while the maximum-likelihood estimator and the Half-Weight Index both estimate *p* correctly (0·50). The latter finding emphasizes the importance of giving careful thought to possible biases caused by different sampling techniques. We have shown that the Half-Weight Index will be most accurate when pairs are more likely to be located when separate than when together (Fig. 2). There is no explicit statement about sampling bias in the model's assumption that *j* out of *k* groups are located at random, and at first sight it seems surprising that the Half-Weight Index estimates *p* accurately under this model. It can be recognized, however, that when sampling a constant proportion of groups, at least one member of a pair will be located more often when the members are separate than when they are together, since if either one of two groups is located the pair can be scored as separate, whereas only one group can provide the information that they are together. In fact, when the formulae for the expected values of different types of sightings are substituted into the formula for the Half-Weight Index, it simplifies to *p*. (This does not, however, mean that the expected value of the Half-Weight Index is *p*, since the expected value of a quotient is not equal to the quotient of the expected values. It does mean that *p* is the value of the first part of the Taylor Series expansion that represents the expected value of the Half-Weight Index, so this expected value is approximately equal to *p*.) Thus, the sampling bias introduced by sampling a constant proportion of existing groups is exactly that needed to obtain an accurate estimate of *p* using the Half-Weight Index when each component of the index takes on its expected value.

The maximum-likelihood estimator and Half-Weight Index both provide accurate estimates of *p* in the absence of random variation. This accuracy is a desirable property of an index, but is by no
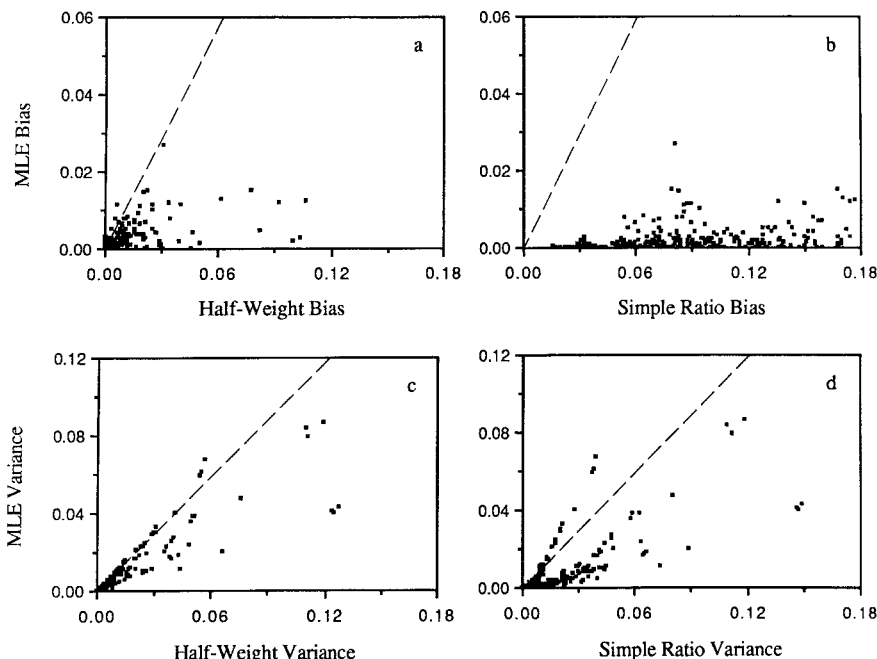
Figure 3. Comparison of bias and variance of the maximum-likelihood estimator (MLE) to those of the Half-Weight and Simple Ratio indices for Model 1. Bias is shown as the absolute difference between the index and the actual proportion of time together used in the simulation. Each plot includes 330 data points, each point representing the mean of 1000 simulated indices. Data points differ in the total number of samples, true association level, number of groups located, and number of groups present, used in the simulation (see Appendix II for actual values). All points below the diagonal line are cases where the Half-Weight or Simple Ratio Index was less accurate than the maximum-likelihood estimator.

means the only criterion for how good an estimate is. The quality of an estimate of p (the probability that two individuals will be together) is determined by the estimate's bias and variance when random variation is present.

A computer simulation was used to compare the variance and bias of the maximum-likelihood estimator under different sampling conditions to those of the Half-Weight and Simple Ratio indices. (The Twice-Weight Index was not included since it had already been shown to be more biased than either the Simple Ratio or Half-Weight Index when the sampling bias is in favour of locating the pair when in separate groups, as is true for this model.) The simulation was similar to the one described earlier. The probabilities of observing a pair together, separately, or not at all, were calculated for given values of j and k (using the formulae in Appendix I). The pair was considered to have been together during the proportion p of the total time periods 'sampled'. Thus, the total number of time periods spent together was calculated as p multi-

plied by the total number of samples (rounded up to the nearest integer). For each of these time periods, a random number was calculated and its value relative to the probability of each type of sighting was used to determine whether x, y_a, y_b, y_{ab} or z would be incremented. The simulation was repeated 1000 times, with all three indices being calculated for each repetition. (The process is similar to flipping a fair coin 10 times, recording the number of heads, and repeating the operation 1000 times. For coin tosses, mathematical proofs have shown that the results should be binomially distributed with a mean of 5 heads in 10 tosses (no bias) and a variance of 2·5. For the association indices, the simulation provides an estimate of bias and variance.) A total of 330 simulations, each consisting of 1000 estimates of the three indices, were run for different combinations of j (ranging from 1 to 18), k (from 3 to 20), p (0·20, 0·50, 0·80) and total days sampled (from 10 to 80). (See Appendix II for actual values.)

Figure 3 compares the variance and bias of the

maximum-likelihood estimator to those of the Half-Weight and Simple Ratio indices for all 330 simulations. Each point represents the average bias or variance observed in 1000 estimates of an index. The absolute value of the average bias of the maximum-likelihood estimator ($\hat{p} - p$) was always less than that of either the Half-Weight or the Simple Ratio, and was low even when their values were high (Fig. 3a, b). Similarly, there were a few cases where the Half-Weight or Simple Ratio had a smaller variance than the maximum-likelihood estimator, but for most cases the variance of the maximum-likelihood estimator was markedly lower than that of either of the other indices (Fig. 3c, d).

The mean square error will be used as a summary measure of how accurate each estimator is. Mean square error is the average squared deviation of an estimator from the true value of the parameter it is estimating. It is equal to the sum of the variance and the square of the bias (thus for an unbiased estimator the mean square error is simply its variance). For the 330 combinations of values used in the simulation, the average mean square errors of the maximum-likelihood estimator, Half-Weight Index and Simple Ratio Index were 0·00623, 0·01005 and 0·02354, respectively. To compare these three values, consider the case of an unbiased estimator with a mean value of 0·50. If its mean square error equalled that of the maximum-likelihood estimator, repeated estimates of its value would fall within the limits of 0·35–0·65 on 95% of all occasions. Increasing its mean square error to that of the Half-Weight Index would increase the width of this interval to 0·30–0·70, while a mean square error equal to that of the Simple Ratio would increase it to 0·20–0·80. Thus under the assumptions of this model the maximum-likelihood estimator is a more precise estimator of p than either the Half-Weight or the Simple Ratio Index. Note that these specific values of the mean square error are based on 1000 estimates of each index.

The maximum-likelihood estimator and Half-Weight Index both have very low mean square error if either the sample size is large or a large proportion of existing groups are located. They differ primarily in variance, and this difference is largest (favouring the maximum-likelihood estimator) when the sample size is small (i.e. $n = 10$ versus 20–80), the true association frequency is high ($p = 0·80$ versus 0·20 or 0·50) and the ratio of j to k is low (few of the existing groups are located).

There are two problems in actually using this maximum-likelihood estimator. (1) It is calculated using k, the total number of groups, which is unknown. (It can be known only if all individuals have been located, in which case the three indices are all equal to the Simple Ratio.) (2) The model's assumption that j and k are fixed is unrealistically simple. The importance of the first problem was tested by calculating the maximum-likelihood estimator using erroneous values for k. Sixteen combinations of j (varying from 1 to 10) and k (from 2 to 25) were used to calculate observed values of x, y_a, y_b, y_{ab} and z equal to the model's expected values. The maximum-likelihood estimator was then calculated using k-values ranging from $j+1$ to $k+j$. In general, large errors in k had little effect on the maximum-likelihood estimator (causing an absolute error of less than 0·05), especially for $j > 2$ and $k > 5$. This suggests that a maximum-likelihood estimator based on an estimated value of k could still be more accurate than another index.

In an attempt to eliminate the second problem, a second maximum-likelihood estimator was derived for a model where j was allowed to vary, but it no longer had a closed-form expression and could be found only by iteration. When this was done using the Newton–Raphson method (Chambers 1977), the maximum-likelihood estimator again had a lower mean square error than the Half-Weight Index. The next step would be to develop a model that allows both j and k to vary, but the derivation of the maximum-likelihood estimator for this model is intractable because k and p are not independent (e.g. if $k = 1$ then p must equal 1). If k is not a constant, the likelihood function has no apparent solution.

These problems do not mean that the maximum-likelihood estimate for the 'j out of k groups' model is never a practical estimator. All statistical analyses are based on models that are simplifications of reality. This is feasible because under suitable conditions (e.g. reasonably large sample size) they are known to be good approximations. Similarly, in a species where k is quite large, the assumption that k does not vary may be acceptable. Another example of field situations in which this model would be close to reality would be studies of bat populations that use a limited number of preferred roost sites (e.g. Wilkinson 1985; Williams 1986), when the researcher has identified some but not all of these roosts. By checking each known roost at

regular intervals, '*j* out of *k* groups' are being located, where *j* and *k* are both fixed as long as each roost (including unknown as well as known sites) contains at least one bat.
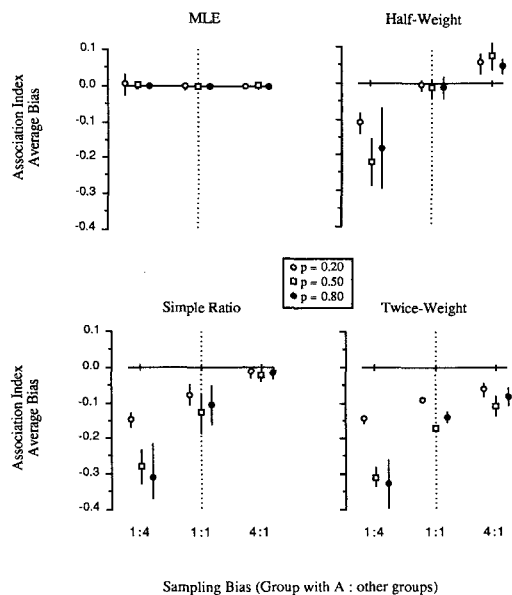

Figure 4. Accuracy of the maximum-likelihood estimator (MLE) for Model II and three other association indices as a function of association frequency (*p*) and sampling conditions. Sampling bias is the relative probability of locating the group containing one member of a pair (A) versus any other group. At the dotted line, the probabilities are equal (no sampling bias). Each point represents the mean (±1 SD) of 60 simulations (each with 1000 repetitions), each one having a different combination of values for the number of groups located, number of groups present and total days sampled (see Appendix II for actual values).

Model II: Differential Observability

One assumption common to all the indices discussed so far is that the two members of a pair are equally likely to be located. It is possible, however, that under certain conditions an individual is less likely to be located than under others. An example of such conditions would include the period of time when a female has young offspring, if she keeps them separate from conspecifics. Another example would be a species where different age–sex classes form separate groups, some of which are less conspicuous than others due to

factors such as the typical size of such groups, their preferred habitats, or the absence of members of a more easily noticed age–sex class (e.g. large or brightly coloured males).

The differential observability model again assumes that *j* out of *k* groups are located, but in addition specifies that one group is observed with a different probability than all other groups. The maximum-likelihood estimator is then derived for a pair, one of whose members (A) is always in that group. (A situation of this type could arise if A was radio-collared, for example: whichever group A was in would be the one with a different probability of being located.) Let *t* equal the observability ratio: if *t* = 2 then the group containing A is twice as likely to be located as any other of the *k* groups. The expected values of *x*, y_a, y_b, y_{ab} and *z* and the derivation of the maximum-likelihood estimator are given in Appendix III.

As with the '*j* out of *k* groups' model, a simulation was used to generate randomly varying values of *x*, y_a, y_b, y_{ab} and *z*. The same combinations of values for *j*, *k* and *p* were used as for the previous model (with a few exceptions: see Appendix II). Only three values for the number of samples were used (*n* = 10, 40 and 80), and the new variable *t* (observability ratio) had the value of either 0·25, 1, or 4. There were thus 540 combinations of values, and again 1000 simulations were run for each combination. For a given combination of *p* and *t* values, there were 60 combinations of values for *j*, *k* and number of samples. The variation over these 60 simulations represents an index's sensitivity to a wide range of possible field conditions. The overall average deviation from the true level of association (±1 SD) for the 60 different sampling conditions is shown in Fig. 4. The maximum-likelihood estimator is the only index that accurately estimates *p* when one group is less likely to be located than the others (*t* = 0·25), and it shows less variation over diverse sampling conditions than any other index for all three sampling biases.

The average mean square errors for the maximum-likelihood estimator, Half-Weight, Simple Ratio and Twice-Weight indices over all 540 combinations were 0·0082, 0·0377, 0·0605 and 0·0785, respectively. Thus the Half-Weight Index again performed better than the Twice-Weight or Simple Ratio Index, but had a mean square error more than four times as large as that of the maximum-likelihood estimator.

DISCUSSION

It should now be clear that the estimation of pairwise association in field studies is a difficult problem. Many variables affect the probability that specific individuals will or will not be located during a given sample. A model that specifies a few assumptions about one or two of these variables may be mathematically tractable but too unrealistic to be of practical use to a field researcher. More detailed models require a larger number of specific assumptions (some of which may turn out to be wrong) and it may not be possible to develop the maximum-likelihood estimator for the parameter of interest. The association indices presently being used appear to be simpler than the maximum-likelihood estimators derived in this study, but are deceptive because they have hidden, implicit assumptions. Comparing their performance to maximum-likelihood estimators under models with explicit assumptions has clarified the differences among them, but their behaviour under conditions other than those examined here is still unknown.

The following conclusions can be drawn.

(1) For any particular study, an index may greatly over- or underestimate the true level of association, depending on how sampling and scoring of association were performed. Comparisons between different studies are therefore difficult unless there is good reason to believe that the indices used were negligibly biased for each study.

(2) A major source of bias in indices of association occurs in sampling situations where the probability of locating a pair depends to some extent on whether or not they are associating. This sampling bias may occur whenever some of the subjects of a study cannot be located at each sample. Each of three currently used indices of association is appropriate for a different degree of sampling bias, as shown in Fig. 2, and a choice between these indices should include an assessment of this bias.

(3) Greater accuracy can be achieved and more powerful analyses performed if an index derived from specific assumptions appropriate for a given sampling situation can be used.

The time to consider this last possibility is before collecting data: it is possible that a little more effort or a slightly different sampling method could make the difference between useful and uninterpretable results. For example, even if it is impossible to locate all individuals on all samples, an especially intensive search made on a previously chosen subset of samples may make it possible to determine what factors make individuals hard to locate. Determining whether a bias exists in the probability of locating a pair when they are together versus separate would allow a decision to be made regarding which would be the best of the currently used indices: the Half-Weight if the bias is in favour of locating the pair when they are in separate groups, the Simple Ratio when no bias exists, or the Twice-Weight when the pair is more likely to be located when together in one group. An even better estimate of association frequency could be made if the critical features of the sampling method were incorporated into a model and the maximum-likelihood estimator derived (as was done for the '*j* out of *k* groups' and the differential observability models). The maximum-likelihood estimator would provide an estimate of association frequency whose properties are well known, and which can be properly evaluated by other researchers.

APPENDIX I

Derivation of the Maximum-likelihood Estimator for Model I

Model I specifies that *j* out of *k* groups are located at each of *n* samples. Assume that all *k* groups are equally likely to be located. The goal is to estimate *p* where *p* is the probability that any pair A, B will be together in the same group at a given moment. Formulae for the probability of each possible type of sighting are given in Table III. The expected values for the number of sightings of each type, based on these probabilities, are given in Table IV. The likelihood function is

$$L\left(p; x, y_a, y_b, y_{ab}, z\right)$$

$$=\binom{n}{x, y_a, y_b, y_{ab}, z}\left[\frac{j}{k}p\right]^x\left[\frac{j(k-j)}{k(k-1)}(1-p)\right]^{y_a+y_b}\left[\frac{j(j-1)}{k(k-1)}(1-p)\right]^{y_{ab}}\left[\frac{(k-j)}{k}p+\frac{(k-j)(k-j-1)}{k(k-1)}(1-p)\right]^z$$

$$=Cp^x(1-p)^{y_a+y_b+y_{ab}}\ (k-j-1+jp)^z$$

Table III. Probabilities of different types of sightings for Model I

True association	Type of sighting	Probability
A, B together	Both A and B	$\binom{k-1}{j-1}\Big/\binom{k}{j} = \dfrac{j}{k}$
A, B together	Neither A nor B	$1-\dfrac{j}{k}=\dfrac{k-j}{k}$
A, B separate	A (or B) only	$\binom{k-2}{j-1}\Big/\binom{k}{j} = \dfrac{j(k-j)}{k(k-1)}$
A, B separate	Both A and B	$\binom{k-2}{j-2}\Big/\binom{k}{j} = \dfrac{j(j-1)}{k(k-1)}$
A, B separate	Neither A nor B	$\binom{k-2}{j}\Big/\binom{k}{j} = \dfrac{(k-j)(k-j-1)}{k(k-1)}$

Table IV. Expected value of the number of sightings of each type under Model I

	Type of sighting	Expected value
x	A, B together	$(j/k)pn$
y_a	A only	$[j(k-j)/k(k-1)](1-p)n$
y_b	B only	$[j(k-j)/k(k-1)](1-p)n$
y_{ab}	A, B separate	$[j(j-1)/k(k-1)](1-p)n$
z	Neither A nor B	$\{[(k-j)/k]p+[(k-j)(k-j-1)/k(k-1)](1-p)\}n$

where the constant C contains all terms not involving p.

The maximum-likelihood estimator is derived as follows

$$\log L = \log C + x \log p + y \log (1-p) + z \log (k-j-1+jp)$$

$$\frac{\mathrm{d}}{\mathrm{d}p}\log L = \frac{x}{p} - \frac{y}{1-p} + \frac{jz}{k-j-1+jp}$$

Setting $\dfrac{\mathrm{d}}{\mathrm{d}p}\log L = 0$ and letting $w = (2j-k+1)x+(j-k+1)y+jz$ gives

$$njp^2 - wp - (k-j-1)x = 0$$

and solving for p gives $\hat{p}$ as defined in the discussion of Model I.

Special cases:

$$\text{Let } z=0; \quad \text{then} \quad p=\frac{x}{x+y}$$

$$\text{Let } j=k-1; \quad \text{then} \quad p=\frac{x+z}{x+y+z}$$

Table V. Expected value of the number of sightings of each type under Model II

	Type of sighting	Expected value
x	A, B together	$f_x rpn$
y_a	A only	$f_{y_a} r(1-p)n$
y_b	B only	$f_{y_b} s(1-p)n$
y_{ab}	A, B separate	$f_{y_{ab}} r(1-p)n$
z	Neither A nor B	$[f_{z_y} s(1-p) + (1-f_x r)p]n$

APPENDIX II

Values used in Simulations of Maximum-likelihood Estimator Models

All combinations of the following values were used in the simulation of Model I. Values in parentheses were excluded from the simulation of Model II in order to keep computing costs feasible.

Total samples: 10, (20), 40, (60), or 80.

True association level (p): 0·20, 0·50, or 0·80.

Number of groups located (j) out of total present (k): for $k=3$, $j=1$ or 2; for $k=5$, $j=2$, (3), or 4; for $k=7$, $j=2, 4$, or 6; for $k=10$, $j=2, 4, 6, 8$, or (9); for $k=15$, $j=3, 6, 9$, or 12; for $k=20$, $j=4, 8, 12, 16$, or 18.

APPENDIX III

Derivation of the Maximum-likelihood Estimator for Model II

Model II specifies that j out of k groups are located at each of n samples. One of the k groups (containing individual A) is observed with a different probability than are all other groups. This differential observability is expressed as a ratio, t, where the probability of observing the group containing A is t times the probability of locating any other group. The goal is to estimate p where p is the probability that any pair A, B will be together in the same group at a given moment. The probability of observing a given set of j groups including the one containing A is

$$r = \sum_{i=1}^{j} \left[\left(\frac{t}{k-i+t} \right)(j-1)! \prod_{l=i}^{(j-1)} \left(\frac{1}{k-l} \right) \prod_{m=1}^{(i-1)} \left(\frac{1}{k-m+t} \right) \right]$$

and the probability of observing a given set of j groups not including the one with A is

$$s = j! \prod_{i=0}^{(j-1)} \left(\frac{1}{k-1+t-i} \right)$$

Let $f_x = \binom{k-1}{j-1}$ be the total possible combinations of j groups which include one containing both A and B when they are together,

$f_{y_a} = f_{y_b} = \binom{k-2}{j-1}$ be the total combinations including only A or B when A and B are separate,

$f_{y_{ab}} = \binom{k-2}{j-2}$ be the total combinations including both A and B when A and B are separate,

and $f_{z_y} = \binom{k-2}{j}$ be the total combinations including neither A nor B when they are separate.

Then the expected values of all types of sighting are as shown in Table V.
The likelihood function is

$$L(p; x, y_a, y_b, y_{ab}, z) =$$

$$M[f_x rp]^x [f_{y_a} r(1-p)]^{y_a} [f_{y_b} s(1-p)]^{y_b} \ [f_{y_{ab}} r(1-p)]^{y_{ab}} [f_{z_y} s(1-p) + (1-f_x r)p]^z$$

where $M = \binom{n}{x, y_a, y_b, y_{ab}, z}$

The derivation of the maximum-likelihood estimator is as follows

$$\log L = \log M + x \log[f_x rp] + y_a \log[f_{y_a} r(1-p)] + y_b \log[f_{y_a} s(1-p)] + y_{ab} \log[f_{y_{ab}} r(1-p)]$$

$$+ z \log[f_{z_y} s(1-p) + (1-f_x r)p]$$

$$\frac{d}{dp}\log L = \frac{x}{p} - \frac{y}{(1-p)} + \frac{hz}{g+hp}$$

where

$$g = f_{z_y} s$$

$$h = 1 - f_x r - f_{z_y} s$$

Set $\dfrac{d}{dp}$ log $L = 0$ and solve for p

$$(1-p)(g+hp)x - p(g+hp)y + p(1-p)hz = 0$$

$$-p^2[hn] + p[h(x+z) - g(x+y)] + gx = 0$$

$$\hat{p} = \frac{-[h(x+z) - g(x+y)] - \sqrt{[h(x+z) - g(x+y)]^2 + 4hgnx}}{-2hn}$$

Special case:

Let $z = 0$; then $\hat{p} = \dfrac{x}{x+y}$

REFERENCES

Altmann, J. 1974. Observational study of behaviour: sampling methods. *Behaviour*, **49**, 227–267.
Chambers, J. M. 1977. *Computational Methods for Data Analysis*. New York: John Wiley.
Clutton-Brock, T.H., Guinness, F.E. & Albon, S.D. 1982. *Red Deer: Behavior and Ecology of Two Sexes*. Chicago: University of Chicago Press.

Dice, L. R. 1945. Amount of ecological association between species. *Ecology*, **26**, 297–302.
Ekman, J. 1979. Coherence, composition, and territories of winter social groups of the willow tit *Parus montanus* and the crested tit *Parus cristatus*. *Ornis Scand.*, **10**, 56–68.
Ficken, M. S., Witkin, S. R. & Weise, C. M. 1981. Associations among individuals in winter flocks of black-capped chickadees. *Behav. Ecol. Sociobiol.*, **8**, 245–249.
Guinness, F. E., Hall, M. J. & Cockerill, R. A. 1979. Mother–offspring association in red deer. *Anim. Behav.*, **27**, 536–544.
Hayaki, H. 1983. The social interactions of juvenile Japanese monkeys on Koshima Islet. *Primates*, **24**, 139–153.
Houston, D. B. 1974. Aspects of the social organization of moose. In: *The Behaviour of Ungulates and its Relation to Management* (Ed. by V. Geist & F.

Walther), pp. 690–696. Morges: International Union for Conservation of Nature and Natural Resources.

Knight, R. R. 1970. The Sun River elk herd. *Wildl. Monogr.*, **23**, 1–66.

Leuthold, B. M. 1979. Social organization and behaviour of giraffe in Tsavo East National Park. *Afr. J. Ecol.*, **17**, 19–34.

Lott, D. F. & Minta, S. C. 1983. Random individual association and social group instability in American bison (*Bison bison*). *Z. Tierpsychol.*, **61**, 153–172.

Mood, A. M., Graybill, F. A. & Boes, D. C. 1974. *Introduction to the Theory of Statistics*. 3rd edn. New York: McGraw-Hill.

Morgan, B. J. T., Simpson, M. J. A., Hanby, J. P. & Hall-Craggs, J. 1976. Visualizing interaction and sequential data in animal behaviour: theory and application of cluster-analysis methods. *Behaviour*, **56**, 1–43.

Myers, J. P. 1983. Space, time and the pattern of individual associations in a group-living species: sanderlings have no friends. *Behav. Ecol. Sociobiol.*, **12**, 129–134.

Penzhorn, B. L. 1984. A long-term study of social organization and behaviour of Cape Mountain zebras *Equus zebra zebra*. *Z. Tierpsychol.*, **64**, 97–146.

Sailer, L. D. & Gaulin, S. J. C. 1984. Proximity, sociality, and observation: the definition of social groups. *Am. Anthropol.*, **86**, 91–98.

Schaller, G. B. 1972. *The Serengeti Lion*. Chicago: University of Chicago Press.

Sorensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. *D. Kgl. Danske Vidensk. Selskab. Biol. Skrifter.*, **5**, 1–34.

Underwood, R. 1981. Companion preference in an eland herd. *Afr. J. Ecol.*, **19**, 341–354.

Wilkinson, G. S. 1985. The social organization of the common vampire bat. I. Pattern and cause of association. *Behav. Ecol. Sociobiol.*, **17**, 111–121.

Williams, C. F. 1986. Social organization of the bat, *Carollia perspicillata* (Chiroptera: Phyllostomidae). *Ethology*, **71**, 265–282.