



FORUM

Testing association patterns of social animals

HAL WHITEHEAD

Department of Biology, Dalhousie University

(Received 9 December 1998; initial acceptance 14 January 1999;
final acceptance 12 February 1999; MS. number: SC-1153)

Indices of association among pairs of animals in a social population naturally vary. An important question in many studies of social structure is whether this variation can be accounted for purely by random association, or whether the data give evidence for preferred association, and/or avoidance of association (Whitehead 1997; Bejder et al. 1998; Whitehead & Dufault 1999). Probably the most appropriate way to carry out such tests is by using Monte-Carlo methods, in which association data are randomly permuted subject to certain constraints, and statistics of the original association data are compared with those from random data. Often the original data set consists of records of which individually identified animals were observed in which group, with group defined as an instantaneously observed spatial aggregation. In these cases it is usual to constrain the permutations so that the number of groups in which each individual was observed, and the number of animals in each group, are held constant (e.g. Slooten et al. 1993; Bejder et al. 1998).

Bejder et al. (1998) permuted association matrices under these constraints by using a sequential routine developed by Manly (1995) for an analogous ecological problem (Table 1). This routine (unlike others that I have used for this task: Whitehead et al. 1982; Slooten et al. 1993) is straightforward, efficient and works with all appropriate data sets. Thus it represents a substantial development in our ability to test social structures for evidence of nonrandom associations.

The core of the method of Manly (1995) and Bejder et al. (1998) (subsequently called MBFB) is the sequential alteration of a 1:0 group–individual matrix in such a way that row and column totals (number of animals in each group, and number of groups of each animal) are unaffected. This is achieved by randomly selecting, at each step, two individuals and two groups so that each individual is seen in only one of the groups, and each

group contains only one of the individuals. The group allocations are then switched. This creates a new matrix with the same row and column totals, which is used as input for the next step (e.g. Table 1).

Test statistics (e.g. median of all half-weight association indices; Bejder et al. 1998) are calculated at each step and used to define the distribution of associations under the null hypothesis, against which the true value of the test statistic can be compared. Test statistics of sequential steps are not independent, but this is not a problem as long as sufficient permutations (steps) are used (Manly 1995). When compared with other procedures, the simplicity and efficiency of Manly's routine greatly outweigh the additional number of permutations that are required.

Here I show that the MBFB method can be extended in two important ways.

Removing Demographic Effects: Permuting Within Sampling Intervals

A statistically significant result from the MBFB procedure indicates that animals associate nonrandomly. This could be because of preferred (or avoided) associations among particular pairs of individuals, that is, a social effect, or it could result from demographic, non-social, factors. For instance, if individuals associate randomly with other members of the population in the study area, but some animals are in the study area for only a portion of the study period (because of habitat selection, death, birth or migration), then associations will appear nonrandom, as individuals whose residence periods were coincident will have unusually high associations, and those pairs that did not co-occur will have a zero association. The MBFB procedure does not distinguish between social and demographic causes of nonrandom association indices.

Demographic effects that result from patterns of residence in the study area can be eliminated as causes for nonrandom association indices by producing random group–individual matrices in which the number of

Correspondence: H. Whitehead, Department of Biology, Dalhousie University, Halifax, Nova Scotia, B3H 4J1 Canada (email: hwhitehe@is.dal.ca).

Table 1. One step in the process of randomizing a group–individual matrix by the method of Manly (1995) and Bejder et al. (1998)

Group	Individual							→	Group	Individual						
	A	B	C	D	E	F	G			A	B	C	D	E	F	G
a	0	1	0	0	0	0	1	a	0	1	0	0	0	0	1	
b	0	1	0	0	0	1	1	b	0	1	0	0	0	1	1	
c	<u>0</u>	1	0	0	0	<u>1</u>	0	c	<u>1</u>	1	0	0	0	<u>0</u>	0	
d	<u>1</u>	0	1	0	0	0	0	d	<u>1</u>	0	1	0	0	0	0	
e	0	0	1	0	1	0	0	e	0	0	1	0	1	0	0	
f	1	0	0	0	1	0	0	f	1	0	0	0	1	0	0	
g	1	0	0	0	1	0	0	g	1	0	0	0	1	0	0	
h	0	0	1	1	0	0	0	h	0	0	1	1	0	0	0	
i	1	0	0	1	0	0	0	i	1	0	0	1	0	0	0	
j	<u>1</u>	0	1	1	0	<u>0</u>	0	j	<u>0</u>	0	1	1	0	<u>1</u>	0	
k	<u>1</u>	0	1	1	1	0	0	k	<u>1</u>	0	1	1	1	0	0	
l	0	0	1	0	1	0	0	l	0	0	1	0	1	0	0	

The matrix on the left, showing which groups contained which individuals, is modified by randomly choosing two individuals and two groups (with each individual in only one of the groups and each group containing only one of the individuals) and switching assignments (underlined), preserving row and column totals.

Table 2. One step in the process of randomizing a group–individual matrix in which ‘between sampling period’ demographic effects are controlled

Interval	Group	Individual							→	Group	Individual						
		A	B	C	D	E	F	G			A	B	C	D	E	F	G
1	a	0	1	0	0	0	0	1	a	0	1	0	0	0	0	1	
1	b	0	1	0	0	0	1	1	b	0	1	0	0	0	1	1	
1	c	0	1	0	0	0	1	0	c	0	1	0	0	0	1	0	
2	d	1	0	1	0	0	0	0	d	1	0	1	0	0	0	0	
2	e	0	0	1	0	1	0	0	e	0	0	1	0	1	0	0	
2	f	1	0	0	0	1	0	0	f	1	0	0	0	1	0	0	
2	g	1	0	0	0	1	0	0	g	1	0	0	0	1	0	0	
3	h	0	0	1	1	0	0	0	h	0	0	1	1	0	0	0	
3	i	<u>1</u>	0	0	1	<u>0</u>	0	0	i	<u>0</u>	0	0	1	<u>1</u>	0	0	
3	j	<u>1</u>	0	1	1	0	0	0	j	<u>1</u>	0	1	1	0	0	0	
3	k	1	0	1	1	1	0	0	k	1	0	1	1	1	0	0	
3	l	<u>0</u>	0	1	0	<u>1</u>	0	0	l	<u>1</u>	0	1	0	<u>0</u>	0	0	

The data were collected in three sampling periods (groups a–c, d–g, h–l). At this step, the third sampling period was randomly chosen, and within it, the matrix was modified by randomly choosing two individuals and two groups and switching assignments (underlined), preserving row and column totals within the sampling period.

groups in which an animal was identified is kept constant within sampling intervals (Whitehead et al. 1982). Sampling intervals should be sufficiently short that movement into, or out of, the study area is unlikely to occur within one sampling interval.

This can be accomplished by modifying the MBFB procedure so that, at each step, a sampling interval is randomly selected first, and then, within that sampling interval, two individuals and two groups are chosen, such that each individual was seen in only one of the groups (Table 2). Otherwise, the analysis proceeds as described by Bejder et al. (1998). This modification can also be used to test whether associations between particular pairs of animals are significantly nonrandom, as in the MBFB procedure.

As an example, the nonzero half-weight association indices calculated from the data in Table 1 (left

side) have a median of 0.00, significantly less than those from random matrices produced using the MBFB procedure (median of 0.40, $P < 0.01$, 5000 permutations). This shows that associations in Table 1 are nonrandom. However, when the sampling intervals are indicated (Table 2, left side), it seems that much of this may be due to just a subset of the animals being present in the study area during each sampling interval. Indeed, using the modified procedure so that column totals are held constant within sampling intervals, there is no longer any statistically significant difference between the mean nonzero half-weight indices in the real and random data (median of random data 0.00, $P = 0.49$, 5000 permutations). Thus the significant nonrandom associations in these data can be fully explained by demographic, nonsocial factors.

Table 3. Randomizing association data

Individual								Individual									
A	B	C	D	E	F	G	H	A	B	C	D	E	F	G	H		
A	—	1	0	0	0	1	0	1	A	—	1	0	1	0	1	0	0
B	1	—	1	1	0	0	1	1	B	1	—	1	1	0	0	1	1
C	0	1	—	1	1	1	1	0	C	0	1	—	0	1	1	1	1
D	<u>0</u>	1	<u>1</u>	—	0	1	0	0	D	<u>1</u>	1	<u>0</u>	—	0	1	0	0
E	0	0	<u>1</u>	0	—	0	1	1	E	0	0	<u>1</u>	0	—	0	1	1
F	1	0	1	1	0	—	0	0	F	1	0	1	1	0	—	0	0
G	0	1	1	0	1	0	—	1	G	0	1	1	0	1	0	—	1
H	<u>1</u>	1	<u>0</u>	0	1	0	1	—	H	<u>0</u>	1	<u>1</u>	0	1	0	1	—

A sampling period is randomly chosen, and the associations within the sampling period, shown in the symmetric 1:0 matrix on the left, are modified by randomly choosing two pairs of individuals (D and H; A and C) so that each individual is associated with only one member of the other pair. The associations between these pairs (underlined) are switched, preserving row (and column) totals.

Association Defined Other Than By Group Membership

In *Bejder et al.'s* (1998) study, animals were considered associated if they were identified within the same group during a sampling period. However, there are other methods of defining association that may be appropriate for different species and populations (*Whitehead & Dufault 1999*). These include distance measures (e.g. 'within x body lengths' or 'nearest neighbours'), temporal measures (e.g. 'dived synchronously'), or behavioural observations (e.g. 'mutual grooming'). If the measure is dichotomous ('not associated'/'associated'), then the data can be expressed as a series of symmetric 0:1 matrices indexed by the study individuals (Table 3), one for each sampling period. These can be used to compute association indices between pairs of animals, for instance using the half-weight index.

When one tests for preferred associations in such cases, an appropriate null hypothesis is that the probability that two individuals are associated in any sampling interval is independent of whether they were associated in other sampling intervals, controlling for the number of associates of each individual in each sampling period. As in the previous modification, this constraint eliminates between-sampling-period demographic effects. A suitable overall test statistic is the standard deviation of the association indices among members of the population: unusually large standard deviations indicate that some animals have preferred (or avoided) companions.

The MBFB method can be modified to produce randomized data sets subject to these constraints, which can be used as the elements for a Monte-Carlo testing procedure. At each step, a sampling interval is chosen randomly, and then two pairs of individuals (making four different individuals) are chosen randomly so that one member of each pair is associated with only one member of the other pair. The associations are then switched (Table 3). The association indices and test statistics are calculated for this new data set, and then it, in turn, is

used as input for the next step, as in the original MBFB method.

Conclusions

With these quite simple extensions, the MBFB method can be used to test a wider range of association patterns, and some demographic causes of nonrandomness in associations can be eliminated. The MBFB technique has the very useful attribute that, in addition to examining the overall pattern of associations in the population, associations between individual diads can be tested against random (*Manly 1995; Bejder et al. 1998*). This ability is conserved in the two modifications discussed above.

It may be possible to extend the MBFB technique in other ways. For instance, association sometimes has continuous rather than 1:0 values (e.g. 'time spent socializing during a sampling period' or 'mean interindividual distance'), and a method of randomizing such data would be useful. Other extensions could include analyses of conditional associations: can the association of A and B be explained simply through their common affiliation with C?

Formal tests of association patterns for departures from randomness add rigour to studies of animal social structure (*Whitehead & Dufault 1999*). I hope that the appearance of the MBFB technique, and the extensions discussed here, will encourage researchers to carry out such tests. The MBFB technique and these extensions are included in a set of programs that I have developed for analysing animal social organization, SOCPROG. Written in the language MATLAB, the programs may be downloaded free from: <http://www.dal.ca/~hwhitehe/social.htm>.

Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council of Canada. Thanks to Michael Appleby, Lars Bejder, Shannon Gowans and an anonymous referee for useful comments on the manuscript.

References

- Bejder, L., Fletcher, D. & Bräger, S. 1998. A method for testing association patterns of social animals. *Animal Behaviour*, **56**, 719–725.
- Manly, B. F. J. 1995. A note on the analysis of species co-occurrences. *Ecology*, **76**, 1109–1115.
- Slooten, E., Dawson, S. M. & Whitehead, H. 1993. Associations among photographically identified Hector's dolphins. *Canadian Journal of Zoology*, **71**, 2311–2318.
- Whitehead, H. 1997. Analysing animal social structure. *Animal Behaviour*, **53**, 1053–1067.
- Whitehead, H. & Dufault, S. 1999. Techniques for analyzing vertebrate social structure using identified individuals: review and recommendations. *Advances in the Study of Behavior*, **28**, 33–74.
- Whitehead, H., Silver, R. & Harcourt, P. 1982. The migration of humpback whales along the northeast coast of Newfoundland. *Canadian Journal of Zoology*, **60**, 2173–2179.