Ecological Monographs, 69(1), 1999, pp. 1–24

DISTANCE-BASED REDUNDANCY ANALYSIS: TESTING MULTISPECIES RESPONSES IN MULTIFACTORIAL ECOLOGICAL EXPERIMENTS

PIERRE LEGENDRE[1] AND MARTI J. ANDERSON[2]

[1]*Département de sciences biologiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada H3C 3J7*
[2]*Centre for Research on Ecological Impacts of Coastal Cities and School of Biological Sciences, Marine Ecology Laboratories, A11, University of Sydney, Sydney, NSW 2006, Australia*

Abstract. We present a new multivariate technique for testing the significance of individual terms in a multifactorial analysis-of-variance model for multispecies response variables. The technique will allow researchers to base analyses on measures of association (distance measures) that are ecologically relevant. In addition, unlike other distance-based hypothesis-testing techniques, this method allows tests of significance of interaction terms in a linear model. The technique uses the existing method of redundancy analysis (RDA) but allows the analysis to be based on Bray-Curtis or other ecologically meaningful measures through the use of principal coordinate analysis (PCoA). Steps in the procedure include: (1) calculating a matrix of distances among replicates using a distance measure of choice (e.g., Bray-Curtis); (2) determining the principal coordinates (including a correction for negative eigenvalues, if necessary), which preserve these distances; (3) creating a matrix of dummy variables corresponding to the design of the experiment (i.e., individual terms in a linear model); (4) analyzing the relationship between the principal coordinates (species data) and the dummy variables (model) using RDA; and (5) implementing a test by permutation for particular statistics corresponding to the particular terms in the model. This method has certain advantages not shared by other multivariate testing procedures. We demonstrate the use of this technique with experimental ecological data from intertidal assemblages and show how the presence of significant multivariate interactions can be interpreted. It is our view that distance-based RDA will be extremely useful to ecologists measuring multispecies responses to structured multifactorial experimental designs.

Key words: experimental ecology; interaction terms in multifactorial ANOVA; intertidal estuarine assemblage; multifactorial experimental design; multivariate permutational tests of interactions; permutations; principal coordinate analysis; redundancy analysis; testing multispecies responses.

INTRODUCTION

The use of structured experiments for testing hypotheses is essential to the science of ecology (Connell 1972, Hurlbert 1984, Underwood 1990, 1997). Ecological experiments often have complex, multifactorial designs. In this context, statistical methods, such as analysis of variance (ANOVA), providing formal tests of hypotheses of individual terms in a complex linear model, are extremely useful.

Tests of interactions among factors form a very important component of multifactorial experiments. Information concerning the interaction of factors is not obtained by single-factor experiments. A significant interaction between two factors indicates that the effects of one of the factors are not consistent across the levels of the other factor. For example, tests of interactions provide a means to test the generality of treatment effects, in space or time (e.g., Hilborn and Stearns 1982, Underwood and Petraitis 1993, Beck 1997). In addition, interactions in multifactorial experiments are the basis for detecting ecological impacts (Green 1979, Underwood 1991, 1992). The individual importance of single factors may be difficult or even impossible to isolate or interpret in a system where interactions are present (Hilborn and Stearns 1982, Underwood 1997).

Some ecological hypotheses make predictions concerning changes of an entire assemblage of species in response to factors. The appropriate analysis to use for tests of interactions with multiple response variables is not as straightforward as applying a linear model to univariate data. Parametric multivariate analysis of variance (MANOVA) requires multivariate normality and homogeneity of covariance matrices. These assumptions are unlikely to be met by many kinds of multivariate ecological data where species are the variables (Legendre and Legendre 1983). Although MANOVA is relatively robust to violations of its assumptions (Olson 1974), the presence of many zeros in a data matrix, caused by rarer species being only occasionally present, will violate the normality assumption, causing problems for these multivariate tests. In addition, the traditional MANOVA statistics are limited by the number of variables they can handle for a given number of replicates. Data sets where there are more

Manuscript received 24 July 1997; revised 18 August 1997; accepted 17 December 1997; final version received 27 April 1998.

species than replicates are common in ecological research. Methodologically, there is a great need for multivariate hypothesis-testing techniques that can be realistically applied in ecology.

In recent years, several nonparametric multivariate statistical methods have emerged (e.g., Mantel 1967, Smith et al. 1990, Clarke 1993). These can be used in an analysis-of-variance approach, but they use permutation tests and thus are free of the assumption of multivariate normality required by parametric MANOVA. Their primary assumption is less stringent, although still important to note: that the data are independently and identically distributed (exchangeability of replicates). Applications of these methods in the ecological literature include (among many others) Hudon and Lamarche (1989), Legendre and Fortin (1989) and Fortin and Gurevitch (1993) (using Mantel's test) and Gray et al. (1988), Anderson and Underwood (1994) and Quinn et al. (1996) (using ANOSIM [Clarke 1993]). These methods are able to test the same kind of null hypothesis as in MANOVA, that is, that assemblages from different treatments (or groups) are no more different than could be expected due to random chance, at a given level of probability.

The statistics used by these methods are based on matrices of distances (or similarities) calculated among replicates. The experimenter has the flexibility to choose a measure of association that is appropriate for the particular data as the basis of the test, rather than being required to rely on a straight-line Euclidean distance as in classical MANOVA. Being based on a distance matrix adds to the usefulness of these nonparametric tests for community ecology, where species may often have nonlinear responses along environmental gradients or to other species or effects (e.g., ter Braak 1987). For many ecological applications, the measure of association used for abundances of species is one that has been proposed independently at least three times—by Steinhaus (in Motyka 1947), Odum (1950), and Bray and Curtis (1957). Here, we refer to this as the ''Bray-Curtis'' measure. It has semi-metric (non-Euclidean) properties but is generally acknowledged to be a good measure of ecological distance for species abundances (Faith et al. 1987, Legendre and Legendre 1998). Alternatives to this coefficient are available in the literature; some, like the coefficient of Kulczynski, are semi-metric, whereas others, like Whittaker's index of association, the Canberra metric, Clark's coefficient of divergence, and the chi-square metric, have metric properties (Legendre and Legendre 1983, 1998).

Despite their appeal and usefulness in ecological studies, these nonparametric methods have the drawback of not generally allowing tests of multivariate interactions between factors in an ANOVA design. In particular, these methods are not designed to differentiate components of multivariate variation in a structured model. The reason for this is twofold. First, the distances used may not be metric, so a linear model cannot be applied directly. The second problem is their reliance on permutation testing. From a strictly mathematical point of view, Edgington (1980) stated that one cannot have an exact permutation test for an interaction term. The reason for this is that the null hypothesis (which generally may be expressed as H_0: $\mu_{11} - \mu_{12} = \mu_{21} - \mu_{22}$) cannot be articulated in terms of the exchangeability of the original replicates. This problem also occurs in permutation tests of interaction for multivariate data carried out using distance-based statistics (Clarke 1993); so it appears that multivariate interaction terms cannot be tested using distance matrices. A test of interaction can only be performed in a model-based context (e.g., ter Braak 1992), using a model that takes into account the main effects. Model-based permutation tests are not strictly exact, however, but only asymptotically exact. Another problem encountered when fitting a model to distance matrices is that many of the distances used by ecologists, such as the Bray-Curtis distance, do not have metric properties; a linear model cannot be directly applied to such distances. This raises the general problem: How can one test for interactions between factors in multivariate data using distance-based tests?

Partitioning of multivariate sums of squares in a multifactorial linear model, which replaces traditional squared straight-line distances with squared dissimilarities, has been described by Pillar and Orlóci (1996). These authors caution, however, that their approach is restricted to the use of dissimilarity measures that have Euclidean metric properties. Unfortunately, many of the dissimilarity measures of greatest relevance to ecologists (including the measure described by Bray and Curtis [1957]), are nonmetric or semi-metric and do not fulfill all of these properties. In addition, for the test of any interaction term in a multifactorial model (with no nested factors included), Pillar and Orlóci (1996) used unrestricted permutation of original replicate vectors, making no mention of the concerns raised by Edgington (1980) on the subject of the null hypothesis.

Here, we describe a new approach, distance-based redundancy analysis (db-RDA). This method is applicable to situations where an ecologist wishes to base the test on the distance measure of his/her own choice, with emphasis on situations where the chosen measure is semi-metric (such as the Bray-Curtis measure), which may lead to non-Euclidean representations. Like ANOSIM or the Mantel test, the db-RDA method is based on a matrix of distances or dissimilarities. The special advantages of db-RDA are that (1) it can be used with distance measures that are non-Euclidean, (2) it can be used to test interaction terms, or any other term in a structured ANOVA model, and (3) it uses nonparametric permutation methods which do not rely on assumptions of multivariate normality.

The purpose of this paper is to describe the db-RDA method (Fig. 1) and to show: (1) the use of principal

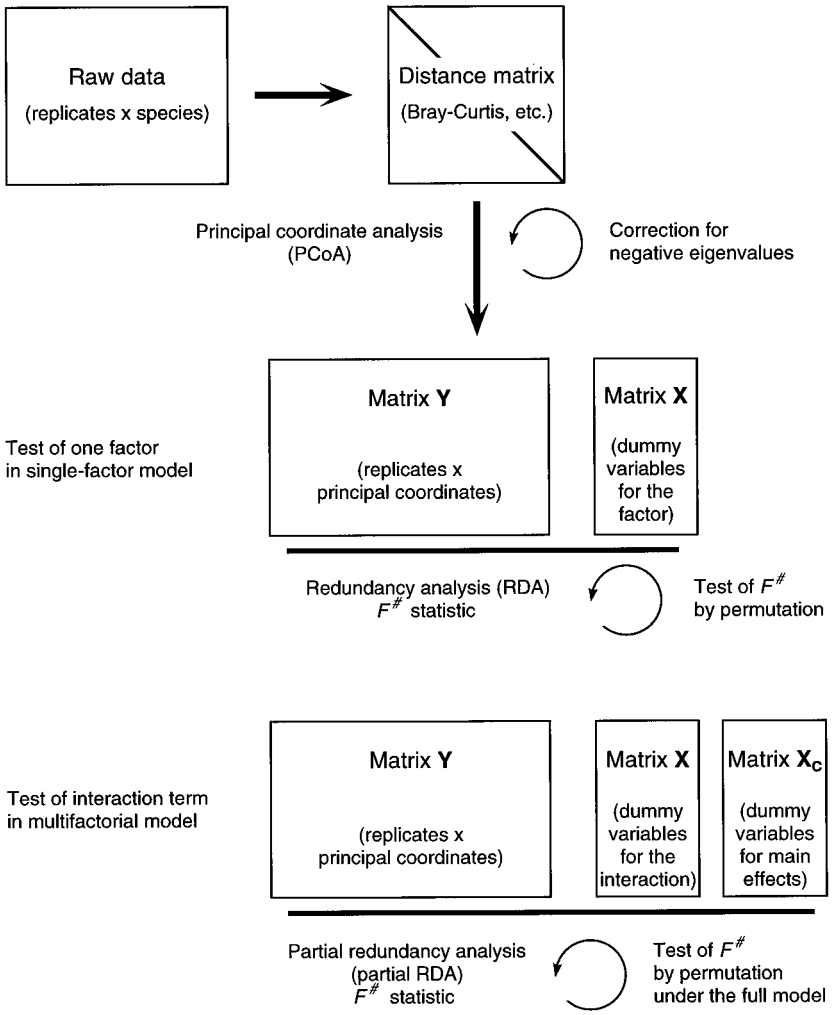

FIG. 1. Graphic outline of the technique for testing multivariate hypotheses in structured experimental designs using distance-based redundancy analysis (db-RDA). Tests of hypotheses concerning effects for single-factor experiments can be done using a single **X** matrix. Tests of interaction terms (or other terms) in a complex multifactorial model require the use of (1) an additional matrix $\mathbf{X}_C$ containing covariables, (2) partial redundancy analysis, and (3) a special method of permutation for tests of individual terms under the full model (see *Description of the technique* for details).

coordinate analysis (PCoA) to place nonmetric or semimetric distances into Euclidean space, so that a linear ANOVA model can be applied; (2) the correspondence between the multivariate RDA statistic and the univariate analysis-of-variance *F* ratio, based on the linear model for multiple regression; and (3) that RDA can be used to test for interactions in multifactorial experimental designs using permutations.

Two examples from experiments on intertidal estuarine assemblages colonizing hard substrata will be used to illustrate the meaning, interpretation, and usefulness of the new technique in ecological applications.

DESCRIPTION OF THE TECHNIQUE

Distance-based redundancy analysis (db-RDA) begins with the calculation of dissimilarities (distances) among replicates, followed by principal coordinate analysis (PCoA), followed by redundancy analysis (RDA), where the **X** matrix (independent variables) contains dummy variables in an ANOVA model and the **Y** matrix (response variables) consists of the principal coordinates. An outline of the method is shown in Fig. 1. The measure of association chosen to calculate distances among replicates in the first instance is of great importance. This choice will greatly influence results and should be justified by the experimenter according to the nature of the data at hand. For a discussion of the many kinds of measures of association that can be used, see Legendre and Legendre (1983, 1998) and Gower and Legendre (1986). We emphasize that the method we propose involving corrections to principal coordinates (in order to place them into a linear Euclidean system) is directed to situations where the ecologist wishes to base the analysis on a non-

Euclidean distance measure. If only the ranks of the distance measure are of interest, or a Euclidean or metric distance is deemed suitable after an appropriate transformation of the data, then other partitioning methodologies can be used (e.g., Clarke 1993, Edgington 1995, Pillar and Orlóci 1996, Manly 1997).

First, we describe the technique of db-RDA, step by step, for one-factor multivariate tests. We then show how it can be used for tests of interactions.

Principal coordinate analysis (PCoA)

Principal coordinate analysis (PCoA, also called ''metric multidimensional scaling'') was described by Gower (1966). PCoA takes a symmetric matrix of distances of any type among replicates and produces corresponding Cartesian (Euclidean) coordinates for each replicate which, in the full-dimensional principal coordinate space, preserve the original distances calculated among replicates (Gower 1966). The procedure of PCoA is briefly summarized here (after Legendre and Legendre 1983, 1998) because intermediate steps will be needed to understand the methods of adjustment for negative eigenvalues (below):

1) Transform the symmetric matrix of distances **D** of elements $\{d_{ij}\}$, $i = 1, \ldots, N$ and $j = 1, \ldots, N$, where N = total number of replicates, into a new matrix **A** of elements $\{a_{ij}\}$ by means of the following equation:

$$a_{ij} = -\frac{1}{2}d_{ij}^2. \quad (1)$$

2) Center the values in matrix **A** by its rows and columns, transforming it into matrix $\mathbf{\Delta}_1$ of elements $\{\delta_{ij}\}$ by means of the following equation:

$$\delta_{ij} = a_{ij} - \bar{a}_i - \bar{a}_j + \bar{a} \quad (2)$$

where $\bar{a}_i$ = average of row i, $\bar{a}_j$ = average of column j and $\bar{a}$ = average of entire matrix **A**.

3) Compute the eigenvalues and eigenvectors of matrix $\mathbf{\Delta}_1$.

4) To obtain principal coordinate axes, scale the eigenvectors to the square root of their respective eigenvalues.

There are, at most, $(N - 1)$ axes required to represent N points in Euclidean space (where N = the total number of replicates). More than one zero eigenvalue may occur, however, so that, in general, the number of axes resulting from a principal coordinate analysis will be less than or equal to $(N - 1)$. The number of principal coordinates depends not only on the number of replicates, but also on the number of variables in the original data matrix **Y** and on the distance measure that has been computed. For the special case of the Euclidean distance, if there are fewer variables (species) than there are replicates in **Y**, then (1) the maximum number of principal coordinates is the number of variables in the original matrix, and (2) the principal coordinates are the same as principal components. Other distance measures may generate more axes than the number of original variables.

For metric distance measures (such as Euclidean or chi-square distances), axes determined using PCoA will preserve all of the original distances, **D**. For nonmetric or semi-metric measures such as Bray-Curtis, the PCoA will only embed the Euclidean part of the distance matrix, with the remainder being given as negative eigenvalues, for which no real axes exist (Gower 1982, 1985). These negative eigenvalues correspond to the variation in the distance matrix that cannot be represented in Euclidean space (Legendre and Legendre 1998). If the principal coordinates corresponding only to the positive eigenvalues were used for the ensuing RDA, the RDA statistic would be a biased estimator of the fraction of the variation of the original data explained by the model matrix, **X**.

Whereas principal coordinate analysis has historically been used for ordination of replicates in a reduced space (i.e., to represent as large a proportion of the variation in the species data as possible, but in a reduced number of dimensions), our use of the method is to retain all of the information in the species data for purposes of hypothesis testing. All axes obtained from these analyses (including after correction for negative eigenvalues, see below) are retained for the test.

Correcting for negative eigenvalues

The original distance matrix can be adjusted to correct for its non-Euclidean portion in the manner shown by Gower and Legendre (1986). Two equally valid yet different solutions exist. We restate here Theorem 7 of Gower and Legendre (1986:10–11), identifying these two methods:

1) A constant, c_1, can be used to correct the squared distances, giving rise to new distances $d'_{ij} = (d_{ij}^2 + 2c_1)^{0.5}$ for $i \neq j$. The value of c_1 is equal to the absolute value of the largest negative eigenvalue of matrix $\mathbf{\Delta}_1$ (Gower and Legendre [1986]; this correction was derived from the earlier work of Lingoes [1971]). The constant c_1 is the smallest value that will produce Euclidean coordinates; any value larger than c_1 will also eliminate all negative eigenvalues and make the system fully Euclidean.

2) A constant, c_2, can be added to all terms d_{ij} of matrix **D** giving rise to new distances $d''_{ij} = d_{ij} + c_2$ for $i \neq j$. The value of c_2 is equal to the largest eigenvalue of the asymmetric matrix:

$$\begin{bmatrix} \mathbf{0} & 2\mathbf{\Delta}_1 \\ -\mathbf{I} & -4\mathbf{\Delta}_2 \end{bmatrix}$$

where $\mathbf{\Delta}_2$ is defined as for $\mathbf{\Delta}_1$ but with elements $-\frac{1}{2}d_{ij}$ rather than $-\frac{1}{2}d_{ij}^2$ (Cailliez 1983, Gower and Legendre 1986). The constant c_2 is the smallest value that will produce Euclidean coordinates; any value larger than c_2 will also eliminate all negative eigenvalues and make the system fully Euclidean (see Fig. 2).

(This description corrects two misprints in Theorem

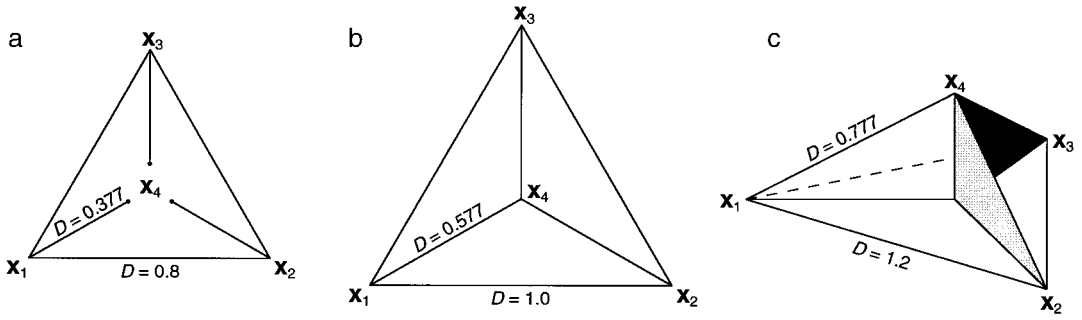

FIG. 2. Distances (D) among four points. (a) Distances are constructed in such a way that the system cannot be represented in Euclidean space because the three lines going toward point $\mathbf{x}_4$ do not meet. (b) By adding a constant to all distances ($c_2 = 0.2$ in the present case), correction method 2 makes the system Euclidean; in this example, the distances can be associated with a representation of the points in two-dimensional space. (c) When increasing the distances further (adding again 0.2 to each distance in the present case), the system remains Euclidean but requires more dimensions for representation (three dimensions in this example).

7 as described by Gower and Legendre [1986]). For each of these correction methods, the values along the diagonal of the matrices **A** or **D** are not changed, but remain as zeros (i.e., the distance between any replicate and itself is zero). Principal coordinate analysis computed on the corrected distance matrix, using either method 1 or method 2, produces a fully Euclidean representation of the replicates without negative eigenvalues or imaginary axes. These axes are then perfectly suitable to apply the linear model of analysis of variance using RDA (or other MANOVA statistic).

TABLE 1. Results of simulations to determine values of correction constants required to correct for negative eigenvalues in principal coordinate analysis for ecological data sets, using the Bray-Curtis measure of dissimilarity. Data were generated randomly using the standardized log-normal distribution, and species weights were generated randomly from a uniform distribution.

No. of simulations	No. of sites	No. of species	Mean values of correction constants	
			Method 1	Method 2
100	30	5	0.16098	0.47448
100	30	10	0.11235	0.37907
100	30	20	0.07682	0.27073
100	30	30	0.05401	0.19160
100	30	50	0.02578	0.09094
50	30	60	0.01440	0.05077
100	30	75	0.00437	0.01535
100	30	100	0.00003	0.00012
50	50	5	0.23441	0.59650
50	50	10	0.16444	0.48815
50	50	20	0.11434	0.38365
100	50	50	0.06225	0.21562
50	50	100	0.01901	0.06602
25	60	3	0.40780	0.73999
50	75	75	0.06535	0.22632
50	75	100	0.04891	0.16925
25	80	4	0.40654	0.79220
25	100	5	0.42778	0.83223
25	100	10	0.25388	0.64015
25	100	20	0.16977	0.51156
25	100	30	0.13871	0.44765
25	100	50	0.10667	0.36096
25	100	75	0.08415	0.28816
25	100	100	0.06842	0.23574
25	100	200	0.02529	0.08677
25	200	10	0.39630	0.82469

As an added note, in some cases (i.e., for certain dissimilarity measures), negative eigenvalues can be eliminated by other kinds of transformations. For example (method 3), all negative eigenvalues are eliminated from a PCoA using the Bray-Curtis measure if the original distances are transformed by taking their square roots (M. J. Anderson, *unpublished simulation results*). Although this result has not yet been proven mathematically, it has been shown that the widely used coefficient of Sørensen, which is the binary form of the Bray-Curtis coefficient, is not Euclidean, while its square root is (Gower and Legendre 1986).

These corrections, although supported by mathematical theorems, may appear quite abstract and even suspect to the practicing ecologist. Four essential questions arise. First, how big are the constants c_1 and c_2 for typical Bray-Curtis distance matrices computed from species abundance data? Second, in what way are the original Bray-Curtis (or other) distances modified by these corrections? Third, what is the influence of making these corrections on the outcome of the $F^{\#}$ test described below? Finally, which of the three possible corrections (in the case of Bray-Curtis distances) should be used? We are most concerned that the thought that goes in to choosing an appropriate ecological distance measure is not sabotaged by any correction, but actually forms the fundamental basis for the ensuing analysis.

We addressed the first question using simulations. Data simulating species abundances were generated randomly using the standardized log-normal distribution. Species weights were obtained randomly from a uniform distribution. After computing Bray-Curtis distances, whose values are between 0 and 1, the corrections (c_1 for method 1, c_2 for method 2) increased nearly linearly with the ratio of the number of replicates to the number of species (Table 1). In extremely species-poor ecosystems, corrections were the largest. For in-

stance, with a ratio of 20:1 (e.g., 200 replicates, 10 species), c_1 was near 0.4 and c_2 was near 0.8. When the ratio was 1:1 (i.e., equal numbers of replicates and species), c_1 was about 0.06 and c_2 was about 0.2. In species-rich ecosystems, corrections were small, becoming smaller as the species richness increased, for a constant number of replicates (Table 1). With a ratio of 1:2 (e.g., 100 replicates, 200 species), c_1 was near 0.02 and c_2 was about 0.1. Results also depended to a small extent on the data-generation parameters.

The second question was addressed directly by comparing (1) the distances in the original Bray-Curtis (or other) distance matrix **D** and (2) the Euclidean distances between replicates described by the new principal coordinate axes. Plots of the relationship between the original Bray-Curtis distances (on the *x*-axis) and the Euclidean distances obtained in the new principal coordinate space (on separate *y*-axes) are shown in Fig. 3 for data on the effects of gastropod grazers at different times (25 species, 72 replicates), discussed in more detail in *Example 1,* below. The top graph shows the relationship when no correction for negative eigenvalues has been done. The imaginary axes corresponding to negative eigenvalues have been left out of the calculations. Obviously the fit is not good; the deviation of points from the ordinary linear regression line indicates the non-Euclidean portion of the Bray-Curtis distances which cannot be embedded directly by the PCoA analysis.

The second plot (middle) shows the relationship when correction method 1 (above) for negative eigenvalues has been used. This correction creates a much smoother line with less scatter, and a concave up relationship with the original distances. A difference in distance at the low end of the scale in the original Bray-Curtis matrix is not as large as the same difference in distance at the high end of the scale. In addition, the entire range of distances has been compressed overall using this method of correction (note the change in scale of the *y*-axis: the coefficient of *x* in the linear equation is 0.4592). The apparent quadratic relationship is not surprising, given the fact that this method creates corrected distances $d'_{ij} = (d_{ij}^2 + 2c_1)^{0.5}$ for $i \neq j$. For these data, the second-order polynomial gives an almost perfect fit ($R^2 = 0.9997$; $y = 0.5525x^2 + 0.0898x + 0.6182$).

When the second method of correction is used ($d''_{ij} = d_{ij} + c_2$ for $i \neq j$), the result is a direct linear relationship between the original Bray-Curtis distances and the Euclidean distances in the corrected principal coordinate space (Fig. 3, bottom, $R^2 = 1.0000$). The only thing that has been done is the addition of a constant to the original distances (except the diagonal); this is also clearly seen in the linear equation for the relationship, showing a slope coefficient of 1 for *x*. By using method 3, taking square roots of Bray-Curtis distances, all distances will be compressed; the relationship with the original distances is simply concave down

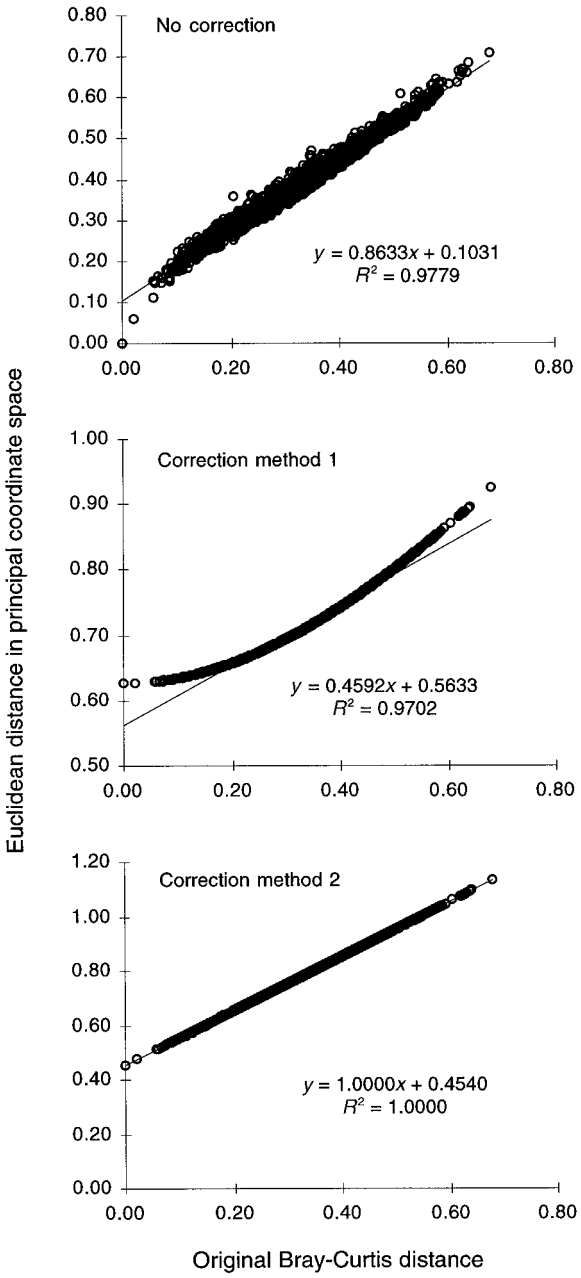

FIG. 3. Plots showing the relationship (ordinary least squares regression) between the original Bray-Curtis distances and the Euclidean distances in principal coordinate space with no correction for negative eigenvalues, using only the axes corresponding to the positive eigenvalues (top), correction of negative eigenvalues using method 1 (middle), or correction of negative eigenvalues using method 2 (bottom) (see *Correcting for negative eigenvalues* for methods). The data are from assemblages on experimental panels in intertidal estuaries in each of three times and three treatments (see *Ecological examples of the technique: Example 1* . . . for details). There were $N = 72$ replicates in the data and thus $N(N - 1)/2 = 2556$ distances among replicates per plot (distances along the diagonal are all equal to zero and were not included).

(dispersion diagram not shown) rather than concave up, as in correction method 1.

These results are not peculiarities of the particular data used, but are general properties of the use of these correction methods for PCoA of dissimilarity matrices (Gower and Legendre 1986).

The third question concerns the effect of correction methods on the $F^{\#}$ test of significance described below (Eqs. 5 and 18). Simulations were used to address this question and the results are contained in Appendix A. We found that, for all methods of correction, the $F^{\#}$ statistic was altered in such a way that, under permutation, values were compressed to be closer to 1.0. That is, relative distances among replicates became less distinct, regardless of the correction method used (Appendix A, Fig. A1).

For correction method 1, in which the change in distances is essentially proportional to their square, all changes to the $F^{\#}$ statistic are monotonic across all permutations (Appendix A, Fig. A2, and Appendix B). Thus, all probabilities, obtained after correction method 1 is applied, are identical to those obtained for tests on the original distances. Consequently, Type I error and power are unchanged by applying correction method 1 (Appendix A, Tables A1 and A2, and Appendix B).

In contrast, the compression of the distribution of the $F^{\#}$ statistic under permutation was not monotonic for correction methods 2 and 3 (Appendix A, Fig. A2). The reference values of $F^{\#}$ appeared too large relative to the compressed permutational distributions. Thus, P values were too small, resulting in inflated Type I error and artificial increases in power with the use of these two methods of correction. These problems were more severe for correction method 3 than for correction method 2, but neither method of correction can be considered appropriate to use with the $F^{\#}$ statistic and test by permutation.

In summary, PCoA, together with the correction for negative eigenvalues, is an integral part of the proposed db-RDA scheme. This procedure allows the calculation of the positions of replicates in Euclidean space, which preserves the original distances computed using any chosen distance measure. This makes the data suitable for analysis using a linear model, such as ANOVA. In answer to our final question, when a correction for negative eigenvalues is needed, method 1 described above is the only suitable method. This correction has no effect on the probabilities for the $F^{\#}$ statistic obtained by permutation. A proof of this theorem is provided in Appendix B.

Redundancy Analysis (RDA)

Redundancy Analysis (RDA) was originally described by Rao (1964, 1973) and popularized by Van den Wollenberg (1977). Referring to redundancy as a quantity, Gittins (1985: 40) defined it as: ''. . . the proportion of the total variance of a measurement domain predictable from a linear composite of the other domain. . . . The term redundancy is therefore synonymous with explained variance.'' RDA has normally been used in ecology for ordination by direct gradient analysis, where a matrix of species variables, **Y** (''measurement domain''), is analyzed with regard to a corresponding matrix of environmental variables, **X** (''other domain'') (e.g., ter Braak 1987, 1990, Roy et al. 1994). This has been the primary use of RDA, which can be done using the computer program CANOCO (ter Braak 1988, 1990). There are, however, examples of the use of RDA in the framework of multivariate hypothesis-testing (Sabatier et al. 1989, ter Braak and Wiertz 1994, Verdonschot and ter Braak 1994). For these applications, instead of a matrix of environmental variables, the ''other domain'' consisted of a matrix of dummy variables describing factors in an experiment. In order to appreciate the use of RDA as a multivariate statistic, it is necessary to understand the multiple regression approach to analysis of variance (see Appendix C).

Canonical eigenvalues and the $F^{\#}$ statistic

We will now show the correspondence between the elements comprising a univariate F statistic in analysis of variance and the multivariate $F^{\#}$ statistic used in redundancy analysis. The correspondence can be understood by reference to (1) the equivalence between univariate analysis of variance and regression using dummy variables (see Appendix C) and (2) an understanding of the algebra of RDA as a series of multiple regressions.

RDA can be described as a series of multiple regressions followed by a principal component analysis (Davies and Tso 1982, ter Braak and Looman 1994, Legendre and Legendre 1998). Here, instead of one vector for one variable, **y**, we are analyzing a data matrix with many variables, **Y**. Each vector variable, **y**, is regressed on the matrix of dummy variables corresponding to the experimental treatments, **X**. The least-squares estimates of the regression coefficients can be found by solving the equation $\mathbf{B} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{Y}$. Using the regression coefficients, one can compute a new matrix, $\hat{\mathbf{Y}}$, of fitted values of the multiple regressions:

$$\hat{\mathbf{Y}} = \mathbf{XB} \quad \text{or} \quad \hat{\mathbf{Y}} = \mathbf{X}[\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{Y}. \tag{3}$$

Computing the principal components of the matrix of fitted values, $\hat{\mathbf{Y}}$, corresponds to solving the following eigenvalue problem:

$$(\mathbf{S}_{\hat{\mathbf{Y}}} - \lambda_h\mathbf{I})\mathbf{u}_h = \mathbf{0} \tag{4}$$

where $\mathbf{S}_{\hat{\mathbf{Y}}}$ is the covariance matrix of the fitted values $\hat{\mathbf{Y}}$. Eigenvalues (λ's) obtained by this expression are the canonical eigenvalues. The number of nonzero canonical eigenvalues (h) is equal to the minimum number of vectors (variables) of either **Y** or **X**.

Note that canonical eigenvalues, as referred to here

and by ter Braak (1987), differ from eigenvalues obtained from unrestricted ordination techniques (such as principal component analysis). Canonical ordination techniques restrict the ordination axes in some way—e.g., to maximize dispersion among groups of replicates (canonical variate analysis), to maximize the relationship between two sets of variables (canonical correlation analysis), to be a linear combination of explanatory variables (RDA), or to be a linear combination of explanatory variables that maximizes dispersion of species scores along a gradient (canonical correspondence analysis) (ter Braak 1987).

The sum of the canonical eigenvalues produced by RDA equals the amount of variation in **Y** explained by the model matrix **X**. Depending on the computer program used, the eigenvalues may be expressed in units corresponding to the total sum of squares in matrix **Y**, the total variation in **Y**, or as a fraction of the total variation in **Y**. The statistic used for testing hypotheses with RDA has been described by ter Braak (1990), as follows:

$$F^{\#} = \frac{\text{trace}/q}{\text{RSS}/(N - q - 1)} \tag{5}$$

where trace = the sum of the canonical eigenvalues as obtained by Eq. 4, RSS = the residual sum of squares (not explained by the canonical eigenvalues), q = the number of variables in matrix **X**. The residual sum of squares is calculated as RSS = (sum of all unconstrained eigenvalues − trace). The sum of all unconstrained eigenvalues is the sum of all of the eigenvalues of the **Y** matrix. This is equivalent to the total sum-of-squares variation in the data. We will use the ''#'' to distinguish the analogous multivariate statistics from the univariate sums of squares, mean squares, or F statistics.

With the use of Monte Carlo permutations to provide P values associated with the $F^{\#}$ statistic, we note that an equivalent statistic is obtained by removing all multiplicative constants, such as degrees of freedom. These remain constant across all permutations for a single data set and thus do not affect the outcome of probabilities obtained (e.g., Edgington 1995, Manly 1997).

In the context of RDA as it is used for direct canonical ordination (''direct gradient analysis'' in ter Braak [1987]), the purpose of the analysis is, for instance, to relate biological species variables (the **Y** matrix) to environmental variables (**X**). In that application, the null hypothesis is that there is no significant relationship between variation in the species data and some linear combination of the environmental variables (i.e., that the environmental variables do not explain a significant proportion of the species data). For present purposes, with structured ecological experiments, the **X** matrix does not contain the values of environmental variables, but rather contains the dummy variables corresponding to the levels of factors in an experiment.

The correspondence between the individual components of the univariate F statistic in ANOVA and components of ter Braak's multivariate $F^{\#}$ statistic, when the **X** matrix contains dummy variables in an experimental design, is shown in Table 2. It is important to note that when RDA is used to analyze results of a structured experiment, as in the present description and application, the number (q) of variables in matrix **X** is equal to the number of linearly independent dummy variables (columns) required for the calculation of the multiple regression (see Appendix C). This is also equal to the number of degrees of freedom for that factor in the model.

The consequence of the correspondence of elements in the construction of the RDA statistic and the univariate F statistic, as shown in Table 2, is that a multivariate statistic ($F^{\#}$), analogous to the univariate F ratio, can be calculated to analyze data for a given term in any ANOVA model, according to the hypothesis being tested. This is illustrated for tests of interaction terms in particular examples below.

The matrix containing the dummy variables coding for the design of the experiment, **X**, does not change in going from the univariate (vector **y**) to the multivariate case (matrix **Y**). The sum of squares for a single factor (i) in the univariate case corresponds to the sum of canonical eigenvalues (trace_i) from the RDA of **Y** on only those variables (columns) in **X** coding for that particular factor (Table 2). The value of q in Table 2 for any factor (i) is the number of degrees of freedom for that factor (df_i) in the model and this will be equal to the number of columns in the **X** matrix used to code for that factor.

Thus, a multivariate analog of the univariate mean square for any factor, A, is

$$\text{MS}_{\text{A}}^{\#} = \frac{\text{trace}_{\text{A}}}{\text{df}_{\text{A}}}. \tag{6}$$

A multivariate analog of the univariate residual sum of squares is

$$\text{SS}_{\text{Res}}^{\#} = (\text{sum of all eigenvalues in } \mathbf{Y}) - \sum_{i=1}^{h} \text{trace}_i \tag{7}$$

where h = the number of terms (i.e., main effects and interactions) in the model. This is equal to the sum of the univariate residual sums of squares across all species. Also, a multivariate analog for the degrees of freedom for the residual in a complex model will be the same as the univariate residual degrees of freedom for that model, namely,

$$\text{df}_{\text{Res}} = N - \left(\sum_{i=1}^{h} \text{df}_{\text{i}}\right) - 1. \tag{8}$$

Thus, a multivariate analog for the residual mean square will be

$$\text{MS}_{\text{Res}}^{\#} = \text{SS}_{\text{Res}}^{\#}/\text{df}_{\text{Res}} \tag{9}$$

and the null hypothesis of no effect of factor A can be

TABLE 2. Correspondence between the various components of the univariate F statistic and the multivariate $F^{\#}$ statistic in the one-factor case. Symbols are as in the text.

Univariate ANOVA		Multivariate RDA statistic
Total sum of squares	$\rightarrow$	sum of all eigenvalues of $\mathbf{Y}$
Treatment sum of squares, SS_{Tr}	$\rightarrow$	trace = sum of canonical eigenvalues of $\mathbf{Y}$ on $\mathbf{X}$
Treatment degrees of freedom, df_{Tr}	$\rightarrow$	q
Residual sum of squares, SS_{Res}	$\rightarrow$	$SS^{\#}_{Res}$ = RSS = sum of all eigenvalues $-$ trace
Residual degrees of freedom = df_{Res}	$\rightarrow$	$df_{Res} = N - q - 1$
Treatment mean square = $SS_{Tr}/df_{Tr} = MS_{Tr}$	$\rightarrow$	$MS^{\#} = \text{trace}/q$
Residual mean square = $SS_{Res}/df_{Res} = MS_{Res}$	$\rightarrow$	$MS^{\#}_{Res} = RSS/(N - q - 1)$
$F = \dfrac{MS_{Tr}}{MS_{Res}}$	$\rightarrow$	$F^{\#} = \dfrac{\text{trace}/q}{RSS/(N - q - 1)}$

Notes: RDA = redundancy analysis, Tr = treatment, $\mathbf{Y}$ = matrix of dependent variables (whose variation is to be explained), $\mathbf{X}$ = matrix of explanatory variables, q = number of variables in matrix $\mathbf{X}$, RSS = residual sum of squares, N = total number of observations.

tested in the multivariate case using the following statistic:

$$F^{\#}_{A} = \frac{MS^{\#}_{A}}{MS^{\#}_{Res}}. \tag{10}$$

The statistic used by RDA carried out on the raw data, with the model matrix composed of dummy variables describing the levels of an experimental factor, is similar to the univariate F statistic. It was called a ''stacked'' F statistic by Verdonschot and ter Braak (1994). We consider it a ''pseudo'' F statistic; it consists of (1) the sums of squares for the factor added up across all species (divided by appropriate df) as its numerator and (2) the residual sums of squares added up across all species (again divided by proper df) as its denominator. The RDA statistic in this form assumes a linear model for the relationships between the species data and the ordination axes (Verdonschot and ter Braak 1994).

In this form, it is difficult to see any real advantage for ecologists of using the regression approach and the RDA statistic over standard MANOVA statistics, apart from the use of permutational testing procedures, which can be used for either method in any event—both use standard linear models and preserve Euclidean distances among replicates. There are, however, certain clear advantages for using the RDA statistic with multivariate ecological data. First of all, if the data matrix $\mathbf{Y}$ is replaced with a matrix of principal coordinates that preserve Bray-Curtis distances (or other ecologically relevant distances) among replicates, then it is clear that an RDA analysis on this matrix will be more appropriate, theoretically, than any linear analysis on the raw data. Corrections of coordinates to eliminate negative eigenvalues using method 1 (described in *Correcting for negative eigenvalues*, above) will have no effect on permutational probabilities, provided the RDA $F^{\#}$ statistic is used. Another argument in favor of RDA is that MANOVA is limited to a number of variables smaller than the total number of replicates, whereas RDA does not suffer such a limitation.

TESTS OF INTERACTIONS USING DISTANCE-BASED RDA (DB-RDA)

There are excellent texts that include full discussions of the construction of F ratios for multifactorial designs in univariate ANOVA (e.g., Snedecor and Cochran 1989, Winer et al. 1991, Sokal and Rohlf 1995, Neter et al. 1996, Underwood 1997). We include here only the information relevant for the explanation of how db-RDA can be used to test multivariate interaction terms. For this, we restrict our discussion to the two-way crossed design. This is done in consideration of what has been stated above and in Appendix C, concerning the relationship between ANOVA and multiple regression using dummy variables, and the information in Table 2.

When an experiment includes two orthogonal (crossed) factors, A and B, and a single response variable, univariate analysis of variance can be used to construct F ratios to test each of three null hypotheses:

H_{01}: There is no interaction between factor A and factor B;

H_{02}: Given that there is independence between any effects of A and B (i.e., no interaction), then there are no differences among levels of factor A;

H_{03}: Given that there is independence between any effects of A and B (i.e., no interaction), then there are no differences among levels of factor B.

The first null hypothesis, stated another way, is that the effect of either factor (if it exists) is not affected by it occurring in different levels of the other factor. Before considering the null hypotheses H_{02} or H_{03}, we must first investigate whether there is any significant interaction between the two factors.

TABLE 3. Two-way crossed ANOVA designs. Symbols are as in the text.

Source	Mean square	Expected mean square†	*F* ratio
a) Two fixed factors			
Among levels of A	MS_A	$\sigma_e^2 + bnK_A^2$	MS_A/MS_{Res}
Among levels of B	MS_B	$\sigma_e^2 + anK_B^2$	MS_B/MS_{Res}
Interaction A × B	MS_{AB}	$\sigma_e^2 + nK_{AB}^2$	MS_{AB}/MS_{Res}
Residual	MS_{Res}	σ_e^2	
b) One fixed, one random factor			
Among levels of A (fixed)	MS_A	$\sigma_e^2 + n\sigma_{AB}^2 + bnK_A^2$	MS_A/MS_{AB}
Among levels of B (random)	MS_B	$\sigma_e^2 + an\sigma_B^2$	MS_B/MS_{Res}
Interaction A × B	MS_{AB}	$\sigma_e^2 + n\sigma_{AB}^2$	MS_{AB}/MS_{Res}
Residual	MS_{Res}	σ_e^2	
c) Two random factors			
Among levels of A	MS_A	$\sigma_e^2 + n\sigma_{AB}^2 + bn\sigma_A^2$	MS_A/MS_{AB}
Among levels of B	MS_B	$\sigma_e^2 + n\sigma_{AB}^2 + an\sigma_B^2$	MS_B/MS_{AB}
Interaction A × B	MS_{AB}	$\sigma_e^2 + n\sigma_{AB}^2$	MS_{AB}/MS_{Res}
Residual	MS_{Res}	σ_e^2	

† In the expected mean square expressions, a = the number of levels in factor A, b = the number of levels in factor B, and n = the number of replicates in each group of the balanced design. K^2 is defined by Eq. 11.

First, we will consider the linear ANOVA model as it applies when the two factors are fixed, then when one factor is fixed and the other is random (a ''mixed model''), and finally when the two factors are random. We define the variance of the main effect of any fixed factor, A, in a univariate analysis as

$$K_A^2 = \frac{\sum_{i=1}^{a} (A_i - \bar{A})^2}{(a - 1)} \tag{11}$$

where $\bar{A}$ is the mean across all levels of factor A, a is the number of levels of factor A, and A_i is the effect of the ith level of factor A. Also, for fixed factors, we assume

$$\sum_{i=1}^{a} A_i = 0. \tag{12}$$

In addition, we define the estimated variance attributable to a random factor, B, as σ_B^2. Similarly, residual error variance is designated by σ_e^2. In general, a fixed factor is a factor for which all of the possible levels of the factor (or at least, all of the possible levels of interest for the study) are included in the experiment. By contrast, a random factor is a factor whose levels are a random subset of all possible levels from a population of levels that could have been included in the study. For a more complete discussion of the distinction between fixed and random factors in biological applications, see Underwood (1981, 1997), Winer et al. (1991), and Sokal and Rohlf (1995).

For the multivariate extension, Eqs. 6–9 concerning calculations of mean squares are true for any term in any ANOVA model. As in the one-way case, the **X** matrix does not change in going from the univariate to the multivariate extension. The dummy variable coding for interaction terms is described in Appendix C. In brief, dummy variables for interaction terms are simply the direct products of the variables coding for the main effects.

Two fixed factors

The expected mean squares and *F* ratios for the model when both factors are fixed are given in Table 3a. For a multivariate analog to the *F* ratio to test for the effect of an interaction term, H_{01}, the RDA statistic is constructed as

$$F_{AB}^{\#} = \frac{MS_{AB}^{\#}}{MS_{Res}^{\#}} \tag{13}$$

where $MS_{AB}^{\#}$ is determined from df_{AB}, which is the number of columns coding for the interaction in matrix **X**, and $SS_{AB}^{\#}$, the sum of canonical eigenvalues of an RDA of **Y** on a subset of matrix **X**, say $\mathbf{X}_{A\times B}$, which includes only those dummy variables coding for the interaction term.

If the interaction term is found to be nonsignificant and the main effects are to be investigated, their corresponding statistics, according to their expected mean squares (Table 3a), are calculated as follows. For a test of the null hypothesis H_{02} above, where there are two fixed factors, the RDA statistic is

$$F_A^{\#} = \frac{MS_A^{\#}}{MS_{Res}^{\#}} \tag{14}$$

and for the test of the null hypothesis H_{03}, the RDA statistic is

$$F_B^{\#} = \frac{MS_B^{\#}}{MS_{Res}^{\#}}. \tag{15}$$

The mixed model: one fixed and one random factor

When the experimental design has one fixed factor (A) and one random factor (B), the calculated mean squares in ANOVA for the sources of variation in the

model are different from the case where there are two fixed factors. These expected mean squares and F ratios are shown in Table 3b. The primary distinction in this model, when compared with the model having two fixed factors, is the change in the expected mean square for factor A when B is random. It contains a component of variation attributable to the interaction term, $n\sigma^2_{AB}$, where n is the number of replicates in each group. For a more complete discussion of mixed-model designs in analysis of variance and pooling procedures when the interaction term can reasonably be considered to be zero, see Underwood (1981, 1997) or Sokal and Rohlf (1995).

For tests of multivariate hypotheses using RDA in such an experiment, all of the codings for dummy variables are exactly the same for random factors and for interaction terms involving fixed and random factors as they were for the case with two fixed factors. In addition, the construction of the $F^{\#}$ statistics to test the interaction term (H_{01}) and the main effect of the random factor B (H_{03}), discussed above, are also exactly the same (Eqs. 13 and 15). The $F^{\#}$ statistic for a test of the main effect A (H_{02}) is, however, not equivalent to Eq. 14 and should be constructed as

$$F^{\#}_{A} = \frac{MS^{\#}_{A}}{MS^{\#}_{AB}} \tag{16}$$

in accordance with the expected mean square for this term in univariate ANOVA shown in Table 3b.

Two random factors

When there are two random factors in an orthogonal design, the resulting expected mean squares and F ratios in the univariate ANOVA for the different terms in the model are shown in Table 3c. In this case, the component of variation attributable to the interaction term in the model, $n\sigma^2_{AB}$, appears in the expected mean squares for each of the main effects, A or B. Thus, the corresponding $F^{\#}$ statistic for the multivariate analog of H_{02} is the same as in Eq. 16. The $F^{\#}$ statistic to test H_{03}, however, should be constructed as

$$F^{\#}_{B} = \frac{MS^{\#}_{B}}{MS^{\#}_{AB}}. \tag{17}$$

Performing the Calculations

When using RDA for tests of multivariate hypotheses in multifactorial models, care must be taken in specifying the denominator mean square of any test. Any test in a structured experiment must reflect the null hypothesis as it is articulated under the full linear analysis-of-variance model (e.g., see Searle 1971).

In the computer package CANOCO, the value of the sum of canonical eigenvalues given by the program (trace_i) will be correct for any particular factor or term in the model (i), so long as the only variables in the **X** matrix (called the ''environmental'' matrix by the program) are those coding for this term alone. The value of RSS in CANOCO, however, is calculated for RDA by subtracting the sum of the canonical eigenvalues (due to **X**) from the total sum of squares in **Y** (see Table 2). In a two-factor experiment, this will not be equal to the $MS^{\#}_{Res}$ if the only variables in the **X** matrix are those that code for the interaction term; the residual will be calculated without taking into account the other terms in the model. This means the F_{RDA} statistic constructed by the program will not be correct for the test, in circumstances where there are other terms in the model.

In normal multiple-regression problems, the effects of a single variable or group of variables in a linear model can be assessed by partialling out the effects of other variables in the model. This is equivalent to treating these ''partialled-out'' effects as covariables in the analysis. The same is true for the application of multivariate RDA to a linear ANOVA model. Care must be taken, however, to ensure that the calculation of a test statistic for a term in a complex ANOVA model corresponds to what is required given expected mean squares and the logic of the experimental design.

In practice, the CANOCO computer program can be used to create the test statistics in situations where the denominator mean square is the residual (e.g., Eqs. 13, 14, and 15). This is so because the program has been written to take into account any designated covariables in the analysis. The actual equation for F_{RDA} given by ter Braak (1990), which had been simplified in Eq. 3 to demonstrate its correspondence with ANOVA for Table 2, is, in fact,

$$F_{RDA} = \frac{\text{trace}/q}{\text{RSS}/(N - p - q - 1)} \tag{18}$$

where p = the number of covariables in the analysis. For any term in the model that has $MS^{\#}_{Res}$ in the denominator of its $F^{\#}$ statistic (as for the two-way interaction term), CANOCO can be used to calculate the correct value of $F^{\#}$ for that term, provided that: (1) the variables in the **X** matrix corresponding to the term in the model being tested are separated from the variables coding for all other terms in the model, the latter being placed into a separate matrix, $\mathbf{X}_C$; and (2) the RDA is done using **X** as the matrix of ''environmental variables'' and using $\mathbf{X}_C$ as the matrix of ''covariables,'' according to the terminology used by CANOCO. Subtraction of p in Eq. 18 corrects the degrees of freedom for the residual, so $MS^{\#}_{Res}$ and $F^{\#}$ will be correctly calculated.

For example, $F^{\#}_{A}$ in Eq. 14 could be calculated by putting the variables coding for factor A in the matrix **X** and putting the variables coding for factor B and for the interaction term in matrix $\mathbf{X}_C$, to be treated as covariables. The correct value of p in a structured experiment, as in the present application, is calculated in a way similar to that of q. It will be the number of orthogonal vectors in the matrix, $\mathbf{X}_C$, which is equal to the sum of the degrees of freedom for factors being coded by the $\mathbf{X}_C$ matrix. The permutation test in CAN-

OCO, described in the next section, provides the permutational probability associated with the F_{RDA} statistic. To test for factor B or for the interaction A×B, matrices $\mathbf{X}$ and $\mathbf{X}_C$ would have to be constructed in a way similar to what we have just described for factor A and new CANOCO runs performed.

In contrast, for tests where the denominator mean square is not the residual (e.g., Eqs. 16 and 17), one must reconstruct the proper value of $F^{\#}$ from separate analyses that give correct sums of squares (traces) for each of the mean squares required for the test. For example, for a test of the term shown in Eq. 16, one would proceed as follows:

1) Do an RDA of $\mathbf{Y}$ on a matrix, $\mathbf{X}_A$, which contains only the dummy variable vectors coding for factor A in the model. The sum of canonical eigenvalues from this analysis is equal to the sums of squares ($SS_A^{\#}$) attributable to factor A.

2) Do an RDA of $\mathbf{Y}$ on a matrix $\mathbf{X}_{A\times B}$, which contains only the dummy variable vectors coding for the interaction between factor A and factor B. The sum of canonical eigenvalues from this analysis is equal to the sums of squares ($SS_{A\times B}^{\#}$) attributable to the interaction.

3) Calculate the mean squares $MS_A^{\#}$ and $MS_{A\times B}^{\#}$ by dividing each of the above $SS^{\#}$ by their degrees of freedom, respectively. The degrees of freedom will be equal to the number of orthogonal dummy variables used to code for each of these terms.

4) Calculate the $F_A^{\#}$ in Eq. 16.

Although we can compute the correct value of $F_A^{\#}$, this value cannot be tested for significance directly using presently available versions of the CANOCO program. One alternative is to use Edgington's method of restricted randomization (Edgington 1995).

As an added note, with RDA the output of CANOCO scales the results of analyses so that the total sums of squares in $\mathbf{Y}$ is 1.00. To obtain true values of the sums of squares for any factor or term, one must multiply the result by the real value of the total sums of squares in $\mathbf{Y}$, which is also given as output by the program. Due to the fact that the value of $F^{\#}$ is a ratio of two mean squares, this extra multiplication step is not necessary for computing these statistics.

Permutation Techniques

The above description has been concerned with the construction of appropriate multivariate statistics for testing particular null hypotheses in two-way crossed designs. The distribution of any of these statistics, with or without any distributional assumptions of individual variables in $\mathbf{Y}$, is unknown. Thus, tests of significance of these statistics require the use of permutations. Indeed, the use of permutations, where no specific assumptions are made concerning the particular distribution of the individual variables or of the statistics themselves, is a major advantage to the use of this proposed technique. This is especially relevant when a correction has been made to eliminate negative eigenvalues for the PCoA (principal coordinate analysis). The permutations, however, may also be restricted according to the null hypothesis and the particular design of the experiment. The theory of permutation/randomization tests and considerations under various experimental designs and models are given in detail by Edgington (1995) and Manly (1997).

In the case of the two-way orthogonal design, whether with fixed or random factors, there has been, however, some difference in opinion concerning the way in which permutations should be done for tests of interactions. Manly (1997) has suggested that random permutations across all replicates can be done, because the null hypothesis of no significant interaction makes no statement concerning the main effects in the model. Edgington (1980), in contrast, stated that no test of interaction could be done by permuting raw data in this way (see *Introduction,* above).

In consideration of this problem, ter Braak (1992) has constructed a methodology of permutations for testing interaction terms, based on the randomization model proposed by Kempthorne (1952) and similar to the bootstrap Monte Carlo tests proposed by Hall and Titterington (1989). By this method, permutations are done of the residuals of the variables in the $\mathbf{Y}$ matrix, after fitting covariables and variables in the $\mathbf{X}$ matrix. It is stated by ter Braak (1990, 1992) that the advantages of using such a permutation method are that (1) interaction terms can be tested, (2) the correlation structure of the $\mathbf{X}$ and $\mathbf{X}_C$ matrices are unchanged during the permutations, and (3) there is more power for the analysis of permuted residuals under the full model (Hall and Titterington 1989). This method of permutation is the default method used by the program CANOCO (version 3.10 and later).

Empirical results from simulations have shown that ter Braak's method of permutation (1992) generally maintains Type I error at chosen significance levels, and has power under alternative hypotheses that is asymptotically equivalent to other methods for multiple regression (Anderson and Legendre, 1999). Although more simulations are required in the context of ANOVA, ter Braak's method of permutation can test interaction terms and was the method we used here. It is also referred to as ''permutations under the alternative hypothesis'' or ''permutations under the full model'' (ter Braak 1992). For a discussion of alternative permutational procedures, see Anderson and Legendre (1999).

If an interaction is found to be nonsignificant and one wishes to test the significance of main effects, there are two primary alternatives. One approach is to use model-based permutation techniques (e.g., Freedman and Lane 1983, ter Braak 1992, Kennedy 1995); the other is to use design-based permutation techniques (e.g., Clarke 1993, Edgington 1995). The design-based approach permutes original data vectors ($\mathbf{Y}$) and tests the significance of an individual factor (A), indepen-

dently of another factor (B), by restricting the permutations to occur within levels of factor B. The model-based permutation methods, on the other hand, permute residuals and can be calculated using CANOCO for tests of $F^{\#}$ statistics where the denominator mean square is the residual (e.g., Eqs. 13, 14, and 15).

Some simulations have been done to determine the error rates associated with using permutation of original replicate vectors vs. permutation of some form of residuals in ANOVA (e.g., Manly 1997). More research in this area is required and is being done (M. J. Anderson and C. J. F. ter Braak, *unpublished manuscript*). Designed-based permutation methods, where it is possible to apply them, have an associated Type I error that is ensured to be the significance level chosen for the test (e.g., Edgington 1995). Permutations of residuals may not always maintain exact Type I error, but may have greater power and flexibility for testing some terms in complex designs. To date, however, no radical differences have been found in terms of Type I or Type II error using these two approaches in ANOVA for testing main effects (Manly 1997, M. J. Anderson and C. J. F. ter Braak, *unpublished manuscript*).

Ecological Examples of the Technique

We shall outline experiments for which the technique of db-RDA was used for tests of multivariate ecological hypotheses for assemblages colonizing intertidal oyster leases in estuaries in New South Wales, Australia. These assemblages, consisting primarily of oysters, barnacles, polychaetes, algae, and other invertebrates, have been studied in detail and are described elsewhere (Anderson and Underwood 1994, Underwood and Anderson 1994, Anderson 1996*b*). Experiments have been done that were designed to test hypotheses about the effects of gastropod grazers on the recruitment and succession of the invertebrates and algae in these assemblages (Anderson and Underwood 1997). There were two aspects of that research that we use here to illustrate the potential usefulness of db-RDA: (1) a test of the consistency of the multivariate effects of grazers at different times and (2) a test for potential indirect effects of grazers. For a more complete description of these experiments, see Anderson (1996*a*) and Anderson and Underwood (1997).

Example 1: a mixed model—consistency through time

Experiments were done at a commercial oyster farm in Quibray Bay (34°01′ S, 151°11′E), part of Botany Bay, south of Sydney, Australia. Experimental panels made of concrete (10×10 cm^2) were used for sampling recruitment. Four panels were attached to sticks, which were then attached to the existing oyster farmers' timber structures at a tidal level of ~0.5 m (see details in Anderson and Underwood [1997]). The effects of the grazing snails, *Austrocochlea porcata* and *Bembicium auratum*, were tested by three treatments: (1) caged sticks, excluding grazers; (2) open sticks, subject to normal grazing; and (3) cage control sticks, where sticks were caged, but grazers were put inside the cages at their natural densities. This latter treatment was a control for the effects of cages not caused by the removal of snails. There were always at least two sticks (eight panels) per treatment. This experiment was done three times, each over a period of 6 mo: October 1993 to April 1994 (Time 1), January to July 1994 (Time 2), and October 1994 to April 1995 (Time 3).

There were two primary factors in this design: Grazers (fixed, 3 levels: open, caged, or cage controls) and Time (random, 3 levels, crossed with the factor of Grazers). Time was considered a random factor: the three levels were the initial times at which each experiment was initiated and these were a subset of many possible times the experiment could have been done. In particular, we wished to test the hypothesis that, if the grazers were affecting the assemblages (or not), their effects (or lack of effect) were consistent at each of the three times investigated. In other words, we wished to test for the significance of the multivariate interaction term: Grazers $\times$ Time.

For all multivariate analyses described here, the raw data (25 variables altogether, of which 18 were invertebrates and 7 were macroscopic algae) were first transformed to $y' = y^{0.25}$, and distances among replicates were calculated using the Bray-Curtis coefficient, as has been done in previous work on these assemblages (Anderson and Underwood 1994, 1997). A visual assessment of a potential multivariate interaction was achieved using nonmetric multidimensional scaling (nMDS) to create an ordination of the replicates (Kruskal and Wish 1978). To test the multivariate interaction term, we used db-RDA (distance-based redundancy analysis). The CANOCO computer program was used where: $\mathbf{X}$ = matrix of dummy variables coding for the interaction term, $\mathbf{X}_C$ = matrix of dummy variables coding for the main effects, and $\mathbf{Y}$ = matrix of principal coordinates obtained from a principal coordinate analysis (PCoA) of the Bray-Curtis distance matrix and corrected for negative eigenvalues as outlined above (method 1 in *Correcting for negative eigenvalues*).

Four panels from Time 3 in the treatment where panels were caged were missing from the data set because one of the experimental sticks was broken and lost. Principal coordinates were computed for all the nonmissing panels. To create a balanced design, the four missing replicate vectors were each replaced with the mean vector of the principal coordinates of the remaining four replicates in that cell (caged panels, Time 3). By using the average values of existing replicates for the missing replicates there is no effect of these ''dummy'' replicates on the estimated average or variance of that combination of treatments for the F test (Underwood 1997).

Fig. 4 shows the nMDS plot for these data. The positioning of the replicates on the ordination corresponds to the ranked distances among replicates. There ap-

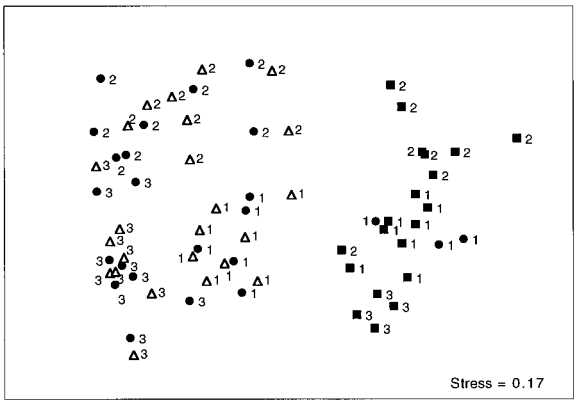

FIG. 4. Two-factor nonmetric multidimensional scaling (nMDS) plot of assemblages on experimental panels in an intertidal estuary showing effects of time and grazing, with each replicate represented by two labels: a number for the time of the experiment (1 = Time 1, 2 = Time 2, and 3 = Time 3) and a symbol for the grazing treatment (△ = open panels; ■ = caged, no grazers; ● = cage controls containing grazers). ''Stress'' is a measure of the disagreement between the inter-point distances on the nMDS plot vs. those distances in the original distance matrix. If stress is relatively low (e.g., < 0.20 [Clarke 1993]), then the plot gives a reasonable representation of the original distances.

peared to be no effect of cages alone in this experiment: open panels (triangles) and cage controls (circles) are quite well mixed in the diagram. The panels that were caged (squares) had quite distinct assemblages, however, forming a distinct separate clump on the plot. Also, the organisms colonizing at different times formed three clear groups: the 2's are located across the top, the 1's in the middle and right, and the 3's at the bottom and left of the plot. Multiple one-way ANOSIM tests (Clarke 1993) comparing the different grazing treatments showed that the three treatments were significantly different from one another at each of the three times (Anderson and Underwood 1997). Although the multiple one-way tests did not suggest any particular interaction between the factors (indeed, a nonsignificant interaction was implied by Anderson and Underwood 1997), there was some variability in the pattern on the nMDS plot. For example, the difference between caged and open (or control) panels was apparently greater during Times 2 and 3 compared to Time 1. A significant multivariate interaction was detected by the db-RDA test ($F^{\#} = 1.358$, $P = 0.001$, 999 permutations).

It was expected that these results could have been obtained (i.e., a significant effect of grazers at each time) purely because grazers were removing algae. Thus, these analyses were redone with algal species omitted. This was to examine the effects of grazers, whether direct or indirect, on other organisms in the assemblage. The nMDS plot showed a similar pattern (see Anderson and Underwood 1997) and the db-RDA was again significant for the interaction term ($F^{\#} = 1.271$ and $P = 0.001$, 999 permutations). This means, ecologically, that the nature of the effects of grazers was somewhat inconsistent through time and was not due purely to their effects on algae alone. This interaction was not readily discernible from multiple one-way tests and thus was not so interpreted by Anderson and Underwood (1997).

Example 2: distinguishing direct from indirect effects

The above results suggested that the effects of grazers on non-algal species, although perhaps variable in time, did occur, but it was not known whether effects were direct (e.g., by grazers ingesting recently settled larvae or by leaving a mucus trail for other organisms to settle on, etc.) or largely indirect, simply due to grazers removing algae and organisms responding in turn to the loss of algae (e.g., by clearing space or by removing a species that facilitated others, etc.). This idea was empirically tested in the third run of the experiment (October 1994 to April 1995). To test for the presence of an indirect vs. a direct effect, the following two factors in a crossed design (as a minimum) are required: (1) Grazers (fixed, 2 levels, present or absent); and (2) Algae (fixed, 2 levels, present or absent). Algae were removed by spraying an industrial herbicide (Reglone [ICI Crop Care, Melbourne, Australia]) on experimental panels, which blocks photosynthesis, affecting microscopic and macroscopic algae, but does not affect other biological tissue.

If there are effects of the presence of grazers that are separate from their effects of removing algae, then there should not be a significant multivariate interaction between these two factors. Alternatively, if grazers are primarily having an indirect effect on assemblages by their removal of algae, then we would expect that assemblages in the following treatments—(+Grazers, +Algae), (+Grazers, −Algae), and (−Grazers, −Algae)—would all be similar and would all differ from assemblages in the treatment of (−Grazers, +Algae). That is to say, in any treatments where grazers are present or algae are removed, the assemblages are the same and these would all be different than where grazers are removed and algae are left to grow. This is one example of how detection of a significant interaction can be ecologically important in a test for indirect effects. The statistical analyses for this experiment were done without including algal species, because these were directly manipulated.

Recall the four missing panels referred to above (see *Example 1*) for Time 3; they are also used in the present example. For the test of this hypothesis, these replicates are now missing from the treatment combination (−Grazers, +Algae). The same strategy was used as described above to replace them for purposes of this analysis.

The nMDS plot of these treatments is shown in Fig. 5. Although each group of replicates corresponding to a particular combination of the two factors forms a rather distinct clump, it is clear that the white triangles (−Grazers, +Algae) separate out on their own much

FIG. 5. Two-factor nonmetric multidimensional scaling (nMDS) plot of assemblages on experimental panels in an intertidal estuary showing effects of removal of grazers and effects of removal of algae. Here, each replicate is identified by two characteristics: (1) squares indicate grazers are present, while triangles indicate grazers are absent; and (2) open symbols indicate algae are present, while solid symbols indicate algae are absent.

further from the other treatments. This is what we would predict if an interaction had occurred. The db-RDA supports this observation, indicating a highly significant multivariate interaction between the two factors ($F^{\#} = 1.831$, $P = 0.001$, 999 permutations). Multiple one-way pairwise ANOSIM analyses further confirmed these results (Anderson 1996*b*). Thus, the biological interpretation of the significant multivariate interaction obtained by db-RDA was, in this case, that the effects of grazers on the assemblage of species were largely indirect, caused by their removal of algae (Anderson 1996*b*).

DISCUSSION

Distance-based RDA, as it is presented here for the first time, offers special advantages to ecological researchers not shared by any other single multivariate method. These are:

1) The researcher has the flexibility to choose an appropriate dissimilarity measure, including those with semi-metric or nonmetric qualities, such as the Bray-Curtis measure.

2) PCoA (principal coordinate analysis) puts the information on dissimilarities among replicates into a Euclidean framework, which can then be assessed using linear models.

3) A correction for negative eigenvalues in the PCoA, if needed, can be done so that probabilities obtained by a permutation test using the RDA $F^{\#}$ statistic are unaffected (correction method 1).

4) By using the multiple-regression approach to analysis of variance, with dummy variables coding for the experimental design, RDA can be used to determine the components of variation attributable to individual factors and interaction terms in a linear model for multivariate data.

5) Multivariate test statistics for any term in a linear model can be calculated, with regard to analogous univariate expected mean squares.

6) Statistical tests of multivariate hypotheses using RDA statistics are based on permutations, meaning that there is no assumption of multinormality of the response variables in the analysis. Also, there are no restrictions to the number of variables that can be included in RDA.

7) Permutations of residuals using the method of ter Braak (1992) allows the permutation test to be structured precisely to the hypothesis and the full linear model of the design under consideration.

8) The significance of multivariate interaction terms can be tested.

Related forms of the kind of technique described here can be found in ter Braak (1992), who described ''constrained principal coordinate analysis,'' and McArdle (1990), who noted that the trace of the regression sum-of-squares matrix can be written in the form of a Mantel statistic.

An alternative to db-RDA is canonical correspondence analysis (CCA; ter Braak 1987). The chi-square distance is the distance preserved among replicates in correspondence analysis (CA) and CCA; it is one of the distances that may be appropriate for species presence–absence or abundance data. If one is satisfied with using this distance measure for a particular application, CCA can be used to test the significance of individual terms in a multifactorial analysis-of-variance model for multispecies data. CCA shares characteristics 4 through 8 with db-RDA. The difference between CCA on the one hand, and db-RDA using the chi-square distance on the other, only resides in the row sums of matrix **Y,** which are used as weights in CCA and not in RDA (ter Braak 1987, Legendre and Legendre 1998).

Traditional parametric methods, such as MANOVA, share characteristics 4, 5, and 8 with the proposed technique. The use of RDA alone shares characteristics 4, 5, 6, 7, and 8 with the proposed method. Other nonparametric techniques based on distance matrices, such as ANOSIM and the Mantel test, share characteristics 1 and 6 with the proposed technique. With the introduction of the use of PCoA and an appropriate correction for negative eigenvalues, distance-based RDA bridges the gap between the traditional approach of linear models in statistics on the one hand, and the need for realistic non-Euclidean measures of association for ecological data on the other.

The use of a correction for negative eigenvalues arising from PCoA on semi-metric distance measures, such as the Bray-Curtis measure, is a unique and valid solution to the problem of applying a linear model to an intrinsically nonlinear ecological system.

Another possible approach, not dissimilar to db-RDA, would be to do MANOVA on the principal coordinates (including correction) obtained from the matrix of distances among replicates. Indeed, the MANOVA statistics of Hotelling's trace or Pillai's trace

should give probabilities equivalent to those obtained using db-RDA (if the coordinates satisfy the distributional assumptions of parametric MANOVA). This is because RDA, as it is used here, can be considered to be a form of straight linear MANOVA: it uses a pseudo *F* statistic consisting of the trace of the among-group sum-of-squares matrix, divided by the trace of the within-group sum-of-squares matrix, multiplied by an appropriate ratio of degrees of freedom. The advantages of using db-RDA for ecological data are its use of permutation testing, compatibility with a correction done on principal coordinates to eliminate negative eigenvalues, leaving probabilities unaltered, and ability to handle large numbers of species variables. The use of Wilks' lambda or other MANOVA statistics may be desirable in certain applications, and permutation-testing procedures can be applied to these statistics (see examples in Manly [1997] and Edgington [1995]). Note, however, that the corrections for negative eigenvalues will have varying and as yet unknown effects on permutation tests with these other multivariate statistics. For some comparisons of various parametric MANOVA statistics and their properties, see Johnson and Field (1993) and Olson (1974, 1976).

The generalized method of db-RDA can be applied to test the significance of individual terms for any complex multifactorial experimental design. Although not outlined here in detail, we suggest that more complex designs will require care with regard to (1) the construction of the multivariate statistic and (2) the permutational procedure used for the test. These should each be tailored to the null hypothesis being tested. Readers are advised to consult Table 2 and Underwood (1981, 1997) for construction of an appropriate $F^{\#}$ statistic (based on univariate ANOVA models) and ter Braak (1992), Edgington (1995), Kennedy (1995), and Manly (1997) concerning permutational procedures for complex experimental designs.

In the form presented here, db-RDA should not be applied to analyze unbalanced data, which are not uncommon for ecological experiments. When the unbalance results from missing or lost values, missing data may be replaced, in principle, by mean vectors, as was done in *Example 2* above. In some situations, however, there may not be enough replicates per cell to justify using vectors of mean coordinates to replace missing replicates. In other instances, whole levels of factors may be missing from a design; this prevents the estimation of a mean response vector.

The method described in this paper for coding linearly independent vectors for factors in the **X** matrix does not result in orthogonal vectors for unbalanced data. Just as standard parametric ANOVA must be modified to accommodate unbalanced designs (e.g., Searle 1987, Shaw and Mitchell-Olds 1993), so would db-RDA. In particular, the dummy variable codings for the **X** matrix would need to be constructed according to Type III or Type IV sums of squares (depending on the nature and degree of missing values). Although not described in detail here, the use of such codings will result in valid multivariate tests using db-RDA for unbalanced data. For a discussion of constructing the **X** matrix for linear models with unbalanced data, see Searle (1987). For a discussion of permutation procedures with unbalanced data, see Edgington (1995).

The method proposed here is not intended to be a panacea for analyzing data from all multivariate experiments. Along with the use of a structured linear model (and the use of ter Braak's permutational strategy) with db-RDA comes the assumption of additivity. This may not be warranted for particular data and is viewed by some as being invalid for a purely nonparametric approach (Edgington 1995).

As an additional note of caution, the db-RDA statistic will be sensitive to heterogeneity of within-group dispersions of the groups. Clarke (1993) discussed this issue with regard to the ANOSIM statistic. Groups may be found to be significantly different on the basis of differences in their dispersions as opposed to differences in their centroids, or central locations. This is the classical Behrens-Fisher problem (Robinson 1982). Traditional parametric MANOVA assumes homogeneity of variance–covariance matrices within groups. Nonparametric statistics using permutation tests do not make this assumption, but in fact the null hypothesis under test with these methods will contain two parts: (1) there are no differences among the central locations (centroids) of the groups *and* (2) there are no differences in the within-group dispersions. Thus, should the null hypothesis be rejected for these nonparametric multivariate tests, there can be no certainty that this is due purely to differences in centroids. A separate nonparametric permutation test for heterogeneity of multivariate dispersions within groups has been developed (M. J. Anderson, P. Dutilleul, F.-J. Lapointe, and P. Legendre, *unpublished manuscript*).

Nevertheless, db-RDA has the potential to become a valuable resource and tool to ecologists who pose multivariate hypotheses and design structured multifactorial experiments to test them. Importantly, it offers a realistic manner whereby the significance of multivariate interaction terms can be tested.

Acknowledgments

Development of this method has benefited from discussions with M. W. Beck, K. R. Clarke, M. G. Chapman, P. Dixon, R. H. Green, B. H. McArdle, C. J. F. ter Braak and A. J. Underwood, as well as from comments from an anonymous reviewer. M. J. Anderson extends particular thanks to B. E. Wingett for his support. This research was supported by an Australian Postgraduate Award and a University of Sydney U2000 Postdoctoral Fellowship to M. J. Anderson and by NSERC grants number OGP7738 and EQP0196089 to P. Legendre.

A FORTRAN program (DISTPCOA: source code and compiled versions for Macintosh and DOS) is available to carry out principal coordinate analysis, including the three corrections for negative eigenvalues described in this paper in ESA's Electronic Data Archive: *Ecological Archives* M069-001.

The program offers the option of calculating the principal coordinates from a pre-computed distance matrix, or computing first one of five distance functions from a table of raw data, with or without preliminary data transformation.

Literature Cited

Anderson, M. J. 1996*a*. Indirect effects of grazing gastropods on recruitment of Sydney rock oysters. Pages 339–347 *in* R. B. Floyd, A. W. Sheppard, and P. J. De Barro, editors. Frontiers of population ecology. Commonwealth Scientific and Industrial Research Organization, Collingwood, Victoria, Australia.

———. 1996*b*. Tests of ecological hypotheses in intertidal estuarine assemblages. Dissertation. University of Sydney, Sydney, New South Wales, Australia.

Anderson, M. J., and P. Legendre. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. Journal of Statistical Computation and Simulation, *in press.*

Anderson, M. J., and A. J. Underwood. 1994. Effects of substratum on the recruitment and development of an intertidal estuarine fouling assemblage. Journal of Experimental Marine Biology and Ecology **184**:217–236.

Anderson, M. J., and A. J. Underwood. 1997. Effects of gastropod grazers on recruitment and succession of an estuarine assemblage: a multivariate and univariate approach. Oecologia **109**:442–453.

Beck, M. W. 1997. Inference and generality in ecology: current problems and an experimental solution. Oikos **78**:265–273.

Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs **27**:325–349.

Cailliez, F. 1983. The analytical solution to the additive constant problem. Psychometrika **48**:305–308.

Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology **18**:117–143.

Connell, J. H. 1972. Community interactions on marine rocky intertidal shores. Annual Review of Ecology and Systematics **3**:169–192.

Davies, P. T., and M. K.-S. Tso. 1982. Procedures for reduced-rank regression. Applied Statistics **31**:244–255.

Draper, N. R., and H. Smith. 1981. Applied regression analysis. John Wiley & Sons, New York, New York, USA.

Edgington, E. S. 1980. Randomization tests. First edition. Marcel Dekker, New York, New York, USA.

———. 1995. Randomization tests. Third edition. Marcel Dekker, New York, New York, USA.

Faith, D. P., P. R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio **69**:57–68.

Fortin, M.-J., and J. Gurevitch. 1993. Mantel tests: spatial structure in field experiments. Pages 342–359 *in* S. M. Scheiner, and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman & Hall, New York, New York, USA.

Freedman, D., and D. Lane. 1983. A nonstochastic interpretation of reported significance levels. Journal of Business and Economic Statistics **1**:292–298.

Gittins, R. 1985. Canonical analysis: a review with applications in ecology. Springer-Verlag, Berlin, Germany.

Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **53**:325–338.

———. 1982. Euclidean distance geometry. Mathematical Scientist **7**:1–14.

———. 1985. Measures of similarity, dissimilarity and distance. Pages 397–405 *in* N. L. Johnson, S. Kotz, and C. B. Read, editors. Encyclopedia of statistics, Volume 5. John Wiley & Sons, New York, New York, USA.

Gower, J. C., and P. Legendre. 1986. Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification **3**:5–48.

Gray, J. S., M. Aschan, M. R. Carr, K. R. Clarke, R. H. Green, T. H. Pearson, R. Rosenberg, and R. M. Warwick. 1988. Analysis of community attributes of the benthic macrofauna of Frierfjord/Langesundfjord and in a mesocosm experiment. Marine Ecology Progress Series **46**:151–165.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, New York, New York, USA.

Hall, P., and D. M. Titterington. 1989. The effect of simulation order on level of accuracy and power of Monte Carlo tests. Journal of the Royal Statistical Society **B 51**:459–467.

Hilborn, R., and S. C. Stearns. 1982. On inference in ecology and evolutionary biology: the problem of multiple causes. Acta Biotheoretica **31**:145–164.

Hudon, C., and G. Lamarche. 1989. Niche segregation between American lobster *Homarus americanus* and rock crab *Cancer irroratus.* Marine Ecology Progress Series **52**:155–168.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs **54**: 187–211.

Johnson, C. R., and C. A. Field. 1993. Using fixed-effects model multivariate analysis of variance in marine biology and ecology. Oceanography and Marine Biology Annual Review **31**:177–221.

Kempthorne, O. 1952. The design and analysis of experiments. John Wiley & Sons, New York, New York, USA.

Kendall, M. G., and A. Stuart. 1963. The advanced theory of statistics. Volume 1. Distribution theory. Second edition. Charles Griffith, London, UK.

Kennedy, P. E. 1995. Randomization tests in econometrics. Journal of Business and Economic Statistics **13**:85–94.

Kruskal, J. B., and M. Wish. 1978. Multidimensional scaling. Sage, Beverly Hills, California, USA.

Legendre, L., and P. Legendre. 1983. Numerical ecology. First English edition. Elsevier, Amsterdam, The Netherlands.

Legendre, P. 1993. Spatial autocorrelation: trouble or a new paradigm? Ecology **74**:1659–1673.

Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio **80**:107–138.

Legendre, P., and Legendre, L. 1998. Numerical ecology. Second English edition. Elsevier, Amsterdam, The Netherlands.

Lingoes, J. C. 1971. Some boundary conditions for a monotone analysis of symmetric matrices. Psychometrika **36**: 195–203.

Manly, B. F. J. 1997. Randomization and Monte Carlo methods in biology. Second edition. Chapman & Hall, London, UK.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research **27**:209–220.

McArdle, B. 1990. Detecting and displaying impacts in biological monitoring: spatial problems and partial solutions. Pages 249–255 *in* Z. Harnos, editor. Proceedings of Invited Papers of the 15th International Biometrics Conference. International Biometric Society, Hungarian Region, Budapest, Hungary.

Motyka, J. 1947. O zadaniach i metodach badan geobotanicznych. Sur les buts et les méthodes des recherches géobotaniques. Annales Universitatis Mariae Curie-Sklodowska , Sectio C, Supplementum I. Lublin, Poland.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. Fourth edition. Irwin, Chicago, Illinois, USA.

Odum, E. P. 1950. Bird populations of the Highlands (North Carolina) plateau in relation to plant succession and avian invasion. Ecology **31**:587–605.

Olson, C. L. 1974. Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association **69**:894–908.
———. 1976. On choosing a test statistic in multivariate analysis of variance. Psychological Bulletin **83**:579–586.
Pillar, V. D. P., and L. Orlóci. 1996. On randomization testing in vegetation science: multifactor comparisons of relevé groups. Journal of Vegetation Science **7**:585–592.
Quinn, G. P., P. S. Lake, and E. S. G. Schreiber. 1996. Littoral benthos of a Victorian Lake and its outlet stream: spatial and temporal variation. Australian Journal of Ecology **21**: 292–301.
Rao, C. R. 1964. The use and interpretation of principal component analysis in applied research. Sankhya **A 26**: 329–358.
———. 1973. Linear statistical inference and its applications. Second edition. John Wiley & Sons, New York, New York, USA.
Robinson, G. K. 1982. Behrens-Fisher problem. Pages 205–209 *in* S. Kotz and N. L. Johnson, editors. Encyclopedia of statistical sciences. Volume 1. John Wiley & Sons, New York, New York, USA.
Roy, R., P. Legendre, R. Knowles, and M. N. Charlton. 1994. Denitrification and methane production in sediment of Hamilton Harbour (Canada). Microbial Ecology **27**:123–141.
Sabatier, R., J.-D. Lebreton, and R. Chessel. 1989. Principal component analysis with instrumental variables as a tool for modelling composition data. Pages 341–352 *in* R. Coppi and S. Bolasso, editors. Multiway data analysis. Elsevier, North Holland, The Netherlands.
Searle, S. R. 1971. Linear models. John Wiley & Sons, New York, New York, USA.
———. 1987. Linear models for unbalanced data. John Wiley & Sons, New York, New York, USA.
Shaw, R. G., and T. Mitchell-Olds. 1993. ANOVA for unbalanced data: an overview. Ecology **74**:1638–1645.
Smith, E. P., K. W. Pontasch, and J. Cairns. 1990. Community similarity and the analysis of multispecies environmental data: a unified statistical approach. Water Research **24**:507–514.
Snedecor, G. W., and W. G. Cochran. 1989. Statistical methods. Eighth edition. Iowa State University Press, Ames, Iowa, USA.
Sokal, R. R., and F. J. Rohlf. 1995. Biometry. Third edition. W. H. Freeman, New York, New York, USA.
ter Braak, C. J. F. 1987. Ordination. Pages 91–169 *in* R. H. Jongman, C. J. F. ter Braak, and O. F. R. van Tongeren, editors. Data analysis in community and landscape ecology. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.
———. 1988. CANOCO: an extension of DECORANA to analyze species–environment relationships. Vegetatio **75**: 159–160.
———. 1990. Update notes: CANOCO version 3.10. Agricultural Mathematics Group, Wageningen, The Netherlands.
———. 1992. Permutation versus bootstrap significance tests in multiple regression and ANOVA. Pages 79–85 *in* K.-H. Jöckel, G. Rothe, and W. Sendler, editors. Bootstrapping and related techniques. Springer-Verlag, Berlin, Germany.
ter Braak, C. J. F., and C. W. N. Looman. 1994. Biplots in reduced-rank regression. Biometrical Journal **36**:983–1003.
ter Braak, C. J. F., and J. Wiertz. 1994. On the statistical analysis of vegetation change: a wetland affected by water extraction and soil acidification. Journal of Vegetation Science **5**:361–372.
Underwood, A. J. 1981. Techniques of analysis of variance in marine biology and ecology. Oceanography and Marine Biology Annual Review **19**:513–605.
———. 1990. Experiments in ecology and management: their logics, functions and interpretations. Australian Journal of Ecology **15**:365–389.
———. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian Journal of Marine and Freshwater Research **42**:569–587.
———. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. Journal of Experimental Marine Biology and Ecology **161**: 145–178.
———. 1997. Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge, England.
Underwood, A. J., and M. J. Anderson. 1994. Seasonal and temporal aspects of recruitment and succession in an intertidal estuarine fouling assemblage. Journal of the Marine Biological Association of the United Kingdom **74**:563–584.
Underwood, A. J., and P. Petraitis. 1993. Structure of intertidal assemblages in different locations: How can local processes be compared? Pages 39–51 *in* R. E. Ricklefs and D. Schulter, editors. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, Illinois, USA.
van den Wollenberg, A. L. 1977. Redundancy analysis: an alternative for canonical correlation analysis. Psychometrika **42**:207–219.
Verdonschot, P. F. M., and C. J. F. ter Braak. 1994. An experimental manipulation of oligochaete communities in mesocosms treated with chlorpyrifos or nutrient additions: multivariate analyses with Monte Carlo permutation tests. Hydrobiologia **278**:251–266.
Winer, B. J., D. R. Broan, and K. M. Michels. 1991. Statistical principles in experimental design. Third edition. McGraw-Hill, Sydney, New South Wales, Australia.

APPENDIX A

Tests of Methods of Correcting for Negative Eigenvalues

Rationale

This Appendix describes results of simulations done to test for the effect of three different methods of transformation to correct for negative eigenvalues (see *Description of the technique: Correcting for negative eigenvalues*) on the $F^{\#}$ statistic and test by permutation. It is not possible to calculate the $F^{\#}$ statistic directly from Bray-Curtis distances in order to compare this with the value of $F^{\#}$ obtained from the principal coordinates after a correction. (If this were possible, principal coordinates would not be necessary in the use of db-RDA.) It is possible, however, to examine the effect of each of the three methods of transformation by comparing the $F^{\#}$ statistic from raw data (calculated using normal Euclidean distances) vs. the $F^{\#}$ statistic as calculated from those data but where a correction method has been used to transform the Euclidean distances to new distances, as follows:

1) Method 1: $d' = (d^2 + 2c_1)^{0.5}$
2) Method 2: $d'' = (d + c_2)$
3) Method 3: $d''' = (d)^{0.5}$.

Methods

Empirical probabilities of Type I error and power were obtained using data simulations. All combinations of the following parameters were used in the generation of data: 1, 5,

TABLE A1. Results of simulations to investigate the effects of three methods of correcting for negative eigenvalues on the Type I error of the multivariate $F^{\#}$ test. Values given are the empirical probabilities of Type I error at the nominal 5% level, calculated from 500 simulations for each combination of parameters; probabilities were obtained using 999 random permutations for each simulation.

No. of variables	No. of groups	No. of replicates per group	No change, d	Method 1† $(d^2 + 2c_1)^{0.5}$	Method 2, $d + c_2$				Method 3, $d^{0.5}$
					$c_2 = 0.1$	$c_2 = 0.2$	$c_2 = 0.5$	$c_2 = 0.9$	
1	2	4	0.052	0.052	0.052	0.052	0.052	0.052	0.052
1	2	7	0.048	0.048	0.050	0.052	0.068	0.076	0.108
1	2	10	0.038	0.038	0.046	0.054	0.066	0.064	0.106
1	5	4	0.088	0.088	0.010	0.106	0.114	0.116	0.132
1	5	7	0.072	0.072	0.092	0.108	0.132	0.156	0.192
1	5	10	0.066	0.066	0.078	0.092	0.144	0.186	0.292
5	2	4	0.042	0.042	0.044	0.044	0.046	0.046	0.058
5	2	7	0.070	0.070	0.076	0.076	0.088	0.100	0.126
5	2	10	0.042	0.042	0.054	0.058	0.070	0.076	0.134
5	5	4	0.076	0.076	0.080	0.086	0.096	0.114	0.142
5	5	7	0.056	0.056	0.072	0.084	0.102	0.122	0.194
5	5	10	0.072	0.072	0.094	0.114	0.164	0.226	0.362
10	2	4	0.022	0.022	0.026	0.026	0.034	0.034	0.034
10	2	7	0.032	0.032	0.030	0.030	0.028	0.026	0.028
10	2	10	0.052	0.052	0.054	0.054	0.056	0.058	0.062
10	5	4	0.058	0,058	0.060	0.062	0.064	0.070	0.078
10	5	7	0.058	0.058	0.058	0.060	0.064	0.068	0.070
10	5	10	0.046	0.046	0.052	0.058	0.068	0.074	0.094

† Regardless of the value of the correction constant, $c_1 = \{0.05, 0.10, 0.25, 0.45\}$, tests on distances after correction using method 1 gave the same probabilities and the same Type 1 errors under permutation as tests of the original distances (d).

or 10 variables; 2 or 5 groups; and 4, 7, or 10 replicates per group. For tests of power, centroids for each group were first generated randomly from a uniform distribution on the interval (0, 1) for each variable. This ensured the null hypothesis of equal centroids was false. These centroids then became the mean vectors for the generation of multivariate normal data. Replicates within each group were generated randomly from a multivariate normal distribution, with all covariance matrices consisting of zero covariances and variances equal to $(\text{SD})^2$. We set the standard deviation of the normal data generator to be SD = 0.2 for data sets with 1 variable, SD = 0.4 for 5 variables, and SD = 0.7 for 10 variables. These values of SD were chosen so that power could be measured (i.e., so that power was >10% and <100%).

For tests of Type I error, data were generated randomly from a multivariate normal distribution, with mean vectors for all groups set at the origin and covariance matrices consisting of zero covariances and variances equal to $(\text{SD})^2$, as described above. For consistency, values of SD were kept the same as indicated above for tests of power, even though this value is of little consequence when the null is true. The same parameters of numbers of groups, variables, and replicates were used for generation of data in obtaining Type I error as were used for investigating power.

After generating data, Euclidean distances were calculated between replicates, then scaled between 0 and 1, so as to conform to the scaling of the Bray-Curtis distance measure. The $F^{\#}$ statistic was calculated on these distances as a basis of comparison with different correction methods. Each of correction methods 1, 2, and 3 was applied, and then the $F^{\#}$ statistic for each was calculated from the corrected distances. A procedure for computing the $F^{\#}$ statistic directly from distances was derived from Theorem 1 of Appendix B. For correction methods 1 and 2 the values of correction constants were chosen given the values of constants found in earlier simulations of ecological data (see Table 1). They were: $c_1 = \{0.05, 0.1, 0.25, 0.45\}$ and $c_2 = \{0.1, 0.2, 0.5, 0.9\}$.

For measuring each of Type I error and power, 500 sets of data were simulated for each set of parameters. For each simulation, $F^{\#}$ statistics were calculated for the uncorrected data and also for the data after each of correction methods 1, 2, and 3 had been applied, and P values were obtained by 999 permutations of the original replicates. Type I error was calculated as the number of rejections of the (true) null hypothesis at $\alpha = 0.05$ (i.e., the number of P values less than 0.05), divided by 500 (the number of simulations). Power was calculated as the number of rejections of the (false) null hypothesis at $\alpha = 0.05$, divided by 500. Type II error is equal to 1 minus the power.

Results

The Type I error was maintained close to nominal α using redundancy analysis (RDA) either with no correction or when correction method 1 was used (Table A1). Type I errors for correction method 1 with increasing values of c_1 are not shown in Table A1, because identical values to those obtained without correction were obtained, regardless of the value of c_1. For correction method 2 or 3, however, the Type I error became inflated with increases in the number of groups (Table A1). Also, the Type I error increased with increases in the value of the correction constant for correction method 2 (Table A1). Correction method 3 had the largest Type I error of all at the nominal 5% level, reaching 36.2% with 5 groups, 5 variables, and 10 replicates.

In addition to the inflation of Type I error, the power of the test was artificially increased by the use of correction method 2 or 3 (Table A2). This gain in power can hardly be viewed as an advantage with these methods of correction, since Type I error is sacrificed when the null hypothesis is true. In contrast, there was no effect on P values and, therefore, there was no effect on probabilities of Type I error or power using correction method 1 (Tables A1 and A2).

Discussion

We briefly outline likely reasons for the above results. Consider the distribution of the $F^{\#}$ statistic under permutation. In

TABLE A2. Results of simulations to investigate the effects of methods of correcting for negative eigenvalues on the power of the multivariate $F^{\#}$ test. Values given are the power (= 1 minus the empirical probability of Type II error at the nominal 5% level) calculated from 500 simulations for each combination of parameters; probabilities were obtained using 999 random permutations for each simulation.

No. of variables	No. of groups	No. of replicates per group	No change, d	Method 1† $(d^2 + 2c_1)^{0.5}$	Method 2, $d + c_2$				Method 3, $d^{0.5}$
					$c_2 = 0.1$	$c_2 = 0.2$	$c_2 = 0.5$	$c_2 = 0.9$	
1	2	4	0.192	0.192	0.192	0.192	0.192	0.192	0.192
1	2	7	0.374	0.374	0.380	0.392	0.406	0.410	0.414
1	2	10	0.482	0.482	0.506	0.518	0.542	0.546	0.574
1	5	4	0.214	0.214	0.236	0.248	0.280	0.294	0.322
1	5	7	0.356	0.356	0.402	0.424	0.486	0.530	0.602
1	5	10	0.504	0.504	0.566	0.610	0.704	0.740	0.796
5	2	4	0.280	0.280	0.296	0.298	0.306	0.310	0.320
5	2	7	0.556	0.556	0.572	0.576	0.602	0.622	0.656
5	2	10	0.732	0.732	0.744	0.760	0.778	0.796	0.834
5	5	4	0.416	0.416	0.438	0.456	0.494	0.516	0.560
5	5	7	0.716	0.716	0.766	0.794	0.824	0.850	0.898
5	5	10	0.852	0.852	0.894	0.912	0.954	0.962	0.978
10	2	4	0.214	0.214	0.216	0.216	0.214	0.210	0.212
10	2	7	0.444	0.444	0.444	0.444	0.454	0.468	0.490
10	2	10	0.672	0.672	0.680	0.684	0.690	0.696	0.714
10	5	4	0.414	0.414	0.420	0.428	0.428	0.440	0.460
10	5	7	0.790	0.790	0.798	0.806	0.816	0.824	0.846
10	5	10	0.944	0.944	0.948	0.952	0.954	0.960	0.970

† Regardless of the value of the correction constant, $c_1 = \{0.05, 0.10, 0.25, 0.45\}$, tests on distances after correction using method 1 gave the same probabilities and the same power under permutation as tests of the original distances (d).

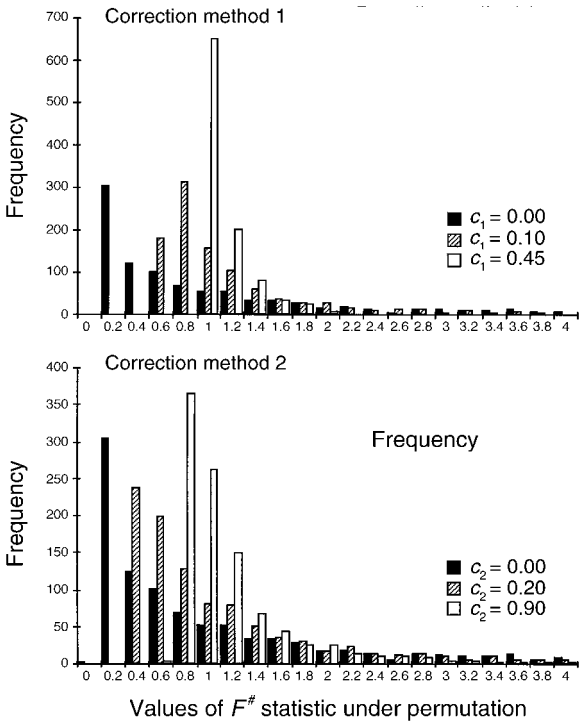

FIG. A1. Frequency distributions of values of the $F^{\#}$ statistic under permutation for one of the data sets generated for measuring power, with one variable, two groups, and 10 replicates per group (see Appendix A: Methods for details of data simulation). The distribution of $F^{\#}$ under permutation is compressed around values near 1.0 with increases in the value of the correction constant for either correction method 1 (top) or method 2 (bottom).

all three methods of correction, the distances become less variable as a consequence of the transformations. For correction method 1, the distances are increased by an amount basically in proportion to their squares (Fig. 3: middle panel). By reference to the original space, although distances are larger in size, they are compressed into a smaller range. For correction method 2, all distances are increased by a constant, meaning the small and large distances are treated equally. Once the new distances are squared in the calculation of sums of squares for the $F^{\#}$ statistic, this correction will have a tendency to decrease the overall separation of points, since the relative increase in small distances (after squaring) is greater than that for large distances. This phenomenon is exacerbated even more with correction method 3, where, by taking square roots of values between 0 and 1.0 (as for the Bray-Curtis measure), all distances are increased in size, but small distances are especially affected, becoming more like the larger distances.

This overall compression of differences between small and large distances is seen in the distribution of the $F^{\#}$ statistic under permutation (Fig. A1). The value of $F^{\#}$ is much less variable after corrections. With increases in the correction constant of either method 1 or method 2, the distribution of values of $F^{\#}$ is condensed into a narrower distribution, closely surrounding the value 1.0 (Fig. A1).

For all methods of correction, the $F^{\#}$ ratio itself changes in value (Table A3). Why is it that the P value for the $F^{\#}$ ratio under permutation is unchanged for method 1, but is decreased for methods 2 and 3 (i.e., resulting in a greater number of rejections: $P \leq 0.05$)?

Recall that the $F^{\#}$ statistic in RDA, like its univariate counterpart, consists of a sum of squared Euclidean distances in each of the numerator and denominator, multiplied by an appropriate ratio of degrees of freedom. The degrees of freedom remain unchanged throughout the permutations, so the ratio of the sums of squares can be considered as an equivalent test statistic. In the case of correction method 1, we note that sums of squares of transformed distances will result in sums of the quantities $(d^2 + 2c_1)$. The numerator of the $F^{\#}$ ratio for the original distances will thus differ from the numerator of the $F^{\#}$ ratio after transformation by the addition of a con-

TABLE A3. $F^{\#}$ statistics and P values associated with a single simulated data set for different methods of correction of negative eigenvalues. The data set had 1 variable, 2 groups, and 10 replicates per group, and the null hypothesis was false.

Method of correction	Value of constant	$F^{\#}$ statistic	P value†
No correction		3.364	0.088
Method 1	$c_1 = 0.05$	2.037	0.088
	$c_1 = 0.10$	1.664	0.088
	$c_1 = 0.25$	1.319	0.088
	$c_1 = 0.45$	1.189	0.088
Method 2	$c_2 = 0.10$	2.992	0.066
	$c_2 = 0.20$	2.649	0.053
	$c_2 = 0.50$	2.030	0.039
	$c_2 = 0.90$	1.670	0.035
Method 3		2.771	0.025

† Using 999 permutations.

stant. Similarly, the denominator of the $F^{\#}$ ratio after transformation using method 1 will also differ from the original only by the addition of a constant. Consequently, although the value of the $F^{\#}$ statistic will change after applying method 1 for correcting negative eigenvalues, the effect on the $F^{\#}$ ratio is monotonic across all permutations, with the constants remaining the same, giving equivalent probabilities for the test (Fig. A2, Table A3). These properties are stated in Appendix B in the form of a theorem.

Therefore, correction method 1 for negative eigenvalues applied to Bray-Curtis distances and used with the $F^{\#}$ statistic in RDA will give the same probabilities under permutation as those that would have been obtained with the original distances (Table A3). For method 2 (and method 3), neither the $F^{\#}$ statistic nor the probabilities under permutation are equivalent to those that would be obtained with the original distances. These transformations are not monotonic with regard to the $F^{\#}$ statistic across all permutations (Fig. A2). P values obtained under permutation are consistently lower than those that would have been obtained with original distances (Table A3).

These simulations show that correction method 1 is the best and only appropriate method to use to correct for negative eigenvalues for permutation tests using the $F^{\#}$ statistic. We note that the probabilities obtained under permutation using correction method 1 will no longer be equivalent to uncorrected data when permuting residuals for the test (e.g., using ter Braak's [1992] method of permutation under the full model for tests of interaction terms). It is already known that, even without any corrections to distances, exact equivalence is not maintained in comparing permutation of raw data, permutation of residuals under the reduced model, or permutation of residuals under the full model for ordinary least-squares regression or ANOVA (e.g., Kennedy 1995, Manly 1997, Anderson and Legendre 1999).

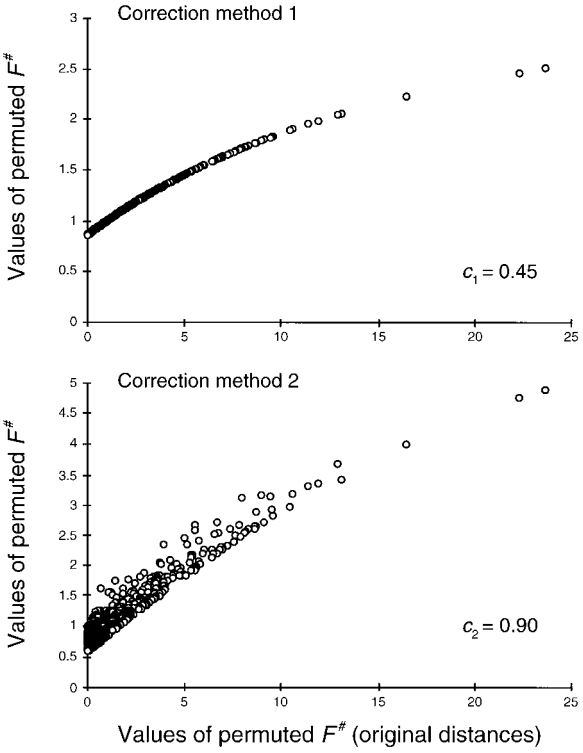

FIG. A2. Values of $F^{\#}$ under permutation for corrected distances vs. those values for original distances, for the same data set as described in Fig. A1. (Top) $F^{\#}$ statistics calculated after using correction method 1 on distances have a clear monotonic relationship with original $F^{\#}$ statistics under permutation. (Bottom) The relationship is not monotonic for correction method 2.

In further simulations we found, however, that when permuting residuals under the full model (ter Braak 1992) as opposed to raw data, the monotonic transformation of the $F^{\#}$ statistic using correction method 1 does not result in any particular bias in Type I error for purposes of the test (Table A4). The parameters used for this simulation study were as outlined above. Results are shown only for the larger sample size ($n = 10$ replicates). Empirical probabilities were much less consistent (with

TABLE A4. Results of simulations to investigate the effects of correction method 1 on the Type I error of the multivariate $F^{\#}$ test when permutation of residuals was used. Values given are the empirical probabilities of Type I error at the nominal 5% level, calculated from 500 simulations for each combination of parameters; probabilities were obtained using 999 random permutations of residuals under the full model for each simulation. There were 10 replicates per group for each data set.

No. of variables	No. of groups	No change, d	Method 1, $(d_2 + 2c_1)^{0.5}$			
			$c_1 = 0.05$	$c_1 = 0.01$	$c_1 = 0.25$	$c_1 = 0.45$
1	2	0.050	0.054	0.058	0.060	0.056
1	5	0.058	0.066	0.068	0.066	0.060
5	2	0.046	0.046	0.046	0.046	0.046
5	5	0.044	0.046	0.046	0.042	0.042
10	2	0.038	0.038	0.038	0.038	0.034
10	5	0.074	0.074	0.072	0.064	0.054

or without the application of a correction) when small sample sizes were used (results not shown). This is a feature of model-based tests. The method of permutation of residuals is not as reliable with very small sample sizes, because the estimates obtained under the model are less accurate with small numbers of replicates (Anderson and Legendre, 1999).

APPENDIX B

Properties of Permutation Tests after Correction Method 1

This Appendix contains a proof showing that, with correction method 1, the addition of a constant, c_1, does not change the test of significance of the analysis-of-variance statistic when using permutation of raw data, although the value of the $F^{\#}$ statistic is changed. The demonstration that follows uses theorem 1, which shows how a multivariate linear sum of squares can be calculated directly from a matrix of Euclidean distances. A proof of this theorem for the univariate case is found in Kendall and Stuart (1963: paragraph 2.22).

Theorem 1

Given matrix $\mathbf{Y} = \{y_{im}\}$, $i = 1, \ldots, n$ and $m = 1, \ldots, p$, a set of n points in p-dimensional Euclidean space. The sum of squares of the distances to the centroid of the group of points is equal to the sum of squared distances among the n points, in the half-matrix of distances $\mathbf{D}$, divided by n. That is,

$$\sum_{m=1}^{p}\sum_{i=1}^{n}(y_{im} - \bar{y}_m)^2 = \frac{1}{n}\sum_{m=1}^{p}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(y_{im} - y_{jm})^2 \quad \text{(B.1)}$$

with

$$\bar{y}_m = \frac{\sum_{i=1}^{n} y_{im}}{n}.$$

Notation.—Let there be g groups and n_k replicates per group ($k = 1, \ldots, g$); n_k are not necessarily equal; $\Sigma\, n_k = N$. Let SS_{Tot0}, SS_{Gr0}, and SS_{Res0} be the total, among-group and between-group (i.e., residual) sums of squares, respectively, for the original distances. Let SS_{Tot2c}, SS_{Gr2c} and SS_{Res2c} be the analogous sums of squares for the distances corrected by addition of a constant $c_1 = c$, in the manner of method 1. The multivariate $F^{\#}$ statistic is: $F_0 = [SS_{Gr0}/(g - 1)]/[SS_{Res0}/(N - g)]$. Values of statistics under permutation will be indicated by an asterisk (e.g., F_0^*).

Lemma.—For corrected distances, sums of squares in an analysis-of-variance framework may be computed using theorem 1. From Eq. B.1, $SS_{Tot0} = (1/N)\,\Sigma\, d_{ij}^2$; therefore, for the corrected distances $d'_{ij} = \sqrt{d_{ij}^2 + 2c}$, we have

$$SS_{Tot2c} = SS_{Tot0} + \frac{[N(N-1)/2] \times 2c}{N} = SS_{Tot0} + (N-1)c \quad \text{(B.2)}$$

and similarly,

$$SS_{Res2c} = SS_{Res0} + \sum_{k=1}^{g}\left\{\frac{[n_k(n_k-1)/2] \times 2c}{n_k}\right\} = SS_{Res0} + (N-g)c \quad \text{(B.3)}$$

$$SS_{Gr2c} = SS_{Tot2c} - SS_{Res2c} = SS_{Gr0} + (g-1)c. \quad \text{(B.4)}$$

Theorem 2

Using permutation of the raw data, the $F^{\#}$ test based upon distances d_{ij} is equivalent to an $F^{\#}$ test based upon distances $d'_{ij} = \sqrt{d_{ij}^2 + 2c}$, which is correction method 1.

Proof.—The $F^{\#}$ statistic, after correction, is

$$F_{2c} = \frac{SS_{Gr2c}/(g-1)}{SS_{Res2c}/(N-g)}. \quad \text{(B.5)}$$

Substituting from Eqs. B.3 and B.4, we have

$$F_{2c} = \frac{[SS_{Gr0} + (g-1)c]/(g-1)}{[SS_{Res0} + (N-g)c]/(N-g)} = \frac{[SS_{Gr0}/(g-1)] + c}{[SS_{Res0}/(N-g)] + c}. \quad \text{(B.6)}$$

It is clear from Eq. B.6 that when $F_0 = 1$, then $F_{2c} = F_0$; when $F_0 < 1$, then $F_0 < F_{2c}$; and when $F_0 > 1$, then $F_0 > F_{2c}$. This explains the compression of permuted values around the value 1 when correction method 1 is used, as observed in Fig. A1 (top).

Now, across all permutations of the raw data (i.e., randomizations of the N replicates), the degrees of freedom remain constant, so F_{2c} is monotonically related to SS_{Gr2c}/SS_{Res2c}. Also, SS_{Tot2c} remains constant across all permutations, therefore SS_{Gr2c} and SS_{Res2c} are monotonically related to one another, varying inversely, and are thus each equivalent test statistics to the F_{2c} statistic under permutation. The same holds for the F_0 statistic without correction; either SS_{Gr0} or SS_{Res0} is an equivalent test statistic to F_0 under raw data permutation, since SS_{Tot0} remains constant. Such equivalent statistics for ANOVA in randomization tests were shown by Edgington (1995).

Importantly, we can see from Eq. B.4 that SS_{Gr0} and SS_{Gr2c} are themselves monotonically related. Thus, $F_{2c}^* > F_{2c}$ if and only if $SS_{Gr2c}^* > SS_{Gr2c}$, which occurs if and only if $SS_{Gr0}^* > SS_{Gr0}$, but $SS_{Gr0}^* > SS_{Gr0}$ if and only if $F_0^* > F_0$; therefore the test of F_{2c} under raw data permutation is equivalent to (i.e., yields the same probability as) the test of F_0.

Note that these strict monotonic relationships between the sums of squares and their associated $F^{\#}$ statistics do not hold if residuals are permuted as opposed to raw data, because SS_{Tot0} (and similarly SS_{Tot2c}) do not stay constant across all permutations. Although not equivalent to permutation of raw data, permutation of residuals is asymptotically unbiased (Freedman and Lane 1983, ter Braak 1992).

APPENDIX C

ANOVA as a Multiple-Regression Linear Model

The first step in the demonstration of the relationship of RDA with ANOVA is to show the equivalence of the ANOVA F ratio with the F ratio produced from a multiple regression of dummy variables corresponding to particular treatments (levels of a factor) in an experiment. There are many statistical texts that describe analysis of variance as a linear model, but more complete descriptions of the specific use of dummy variables in the regression approach to ANOVA are found in Searle (1971), Draper and Smith (1981) and Neter et al. (1996).

In multiple regression, where y_i are $i = 1, \ldots, N$ independent random variables (response), the familiar linear model is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \quad \text{(C.1)}$$

where x_{ij}'s (k predictor variables) are known constants, β_0, $\beta_1, \ldots, \beta_k$ are unknown parameters, and ε_i's are independent normal random variables. In matrix notation, the model is written as $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, with matrices

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ 1 & x_{N1} & \cdot & \cdot & x_{Nk} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_N \end{bmatrix}.$$

The least-squares estimates $\mathbf{b}$ of the parameters β are obtained by the normal equations

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y} \tag{C.2}$$

and taking the inverse of $\mathbf{X}'\mathbf{X}$, we have

$$\mathbf{b} = [\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{y}]. \tag{C.3}$$

In a similar fashion, consider the linear model for a one-way ANOVA:

$$y_{ij} = \mu + t_i + \varepsilon_{ij} \tag{C.4}$$

where y_{ij} is the value of the jth replicate in the ith treatment, μ is the overall parametric mean, t_i is the effect of the ith treatment, and ε_{ij} is the random normal error associated with that replicate. The model for the expectation of y in any particular treatment is

$$E(y_i) = \mu + t_i. \tag{C.5}$$

If there were, for example, three treatments, the model could be written as

$$E(y) = \mu X_0 + t_1X_1 + t_2X_2 + t_3X_3. \tag{C.6}$$

The values of X_i required to reproduce the model $E(y_i) = \mu + t_i$ for a given y_i, using equation C.6, are

$$X_0 = 1$$

$$X_i = \begin{cases} 1 & \text{if the } i\text{th treatment is applied,} \\ 0 & \text{otherwise} \end{cases}.$$

This can be expressed by the following matrices:

$$\mathbf{y} = \begin{bmatrix} y_{11} \\ \cdot \\ \cdot \\ y_{1j} \\ \hline y_{21} \\ \cdot \\ \cdot \\ y_{2j} \\ \hline y_{31} \\ \cdot \\ \cdot \\ y_{3j} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \mu \\ t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

where the columns of the matrix $\mathbf{X}$ correspond to X_0, X_1, X_2, and X_3, respectively. A least-squares solution may again be obtained by the equation

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}. \tag{C.7}$$

The above $\mathbf{X}$ matrix has column dependencies, each binary variable coding for the factor being a linear combination of all the others. This causes $\mathbf{X}'\mathbf{X}$ to be singular, with no inverse and thus no unique solution for $\mathbf{b}$. Thus, the ANOVA model is said to be ''overparameterized.'' What is required is a matrix equivalent to $\mathbf{X}$ but that has linearly independent (orthogonal) columns (see the following section, as well as Draper and Smith [1981]).

The regression sum of squares (and mean square) calculated from the above least-squares solution for $\mathbf{b}$ is exactly the same as the treatment sum of squares (and mean square) that would be calculated in the normal fashion in an ANOVA (Draper and Smith 1981, Legendre 1993). In multiple regression we refer to an ''explained'' sum of squares, or the amount of total variation in $\mathbf{y}$ that can be explained by $\mathbf{X}$. The multiple-regression approach to ANOVA similarly determines the amount of variation in $\mathbf{y}$ explained by the model $\mathbf{X}$ matrix, that is, by our ANOVA model.

Coding of orthogonal dummy variables

A few comments concerning the use of dummy variables are appropriate. When more than one factor is considered in an analysis, it is most efficient to use vectors of dummy variables that are orthogonal. See Draper and Smith (1981) for a discussion of the importance of orthogonality in the use of dummy variables. A general pattern of coding that can be used routinely to obtain orthogonal vectors for balanced experiments (i.e., experiments with the same number of replicates in each group) is succinctly outlined by the following matrices:

2 levels, 1 vector	3 levels, 2 vectors	4 levels, 3 vectors	5 levels, 4 vectors
$\begin{bmatrix} +1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} +2 & 0 \\ -1 & +1 \\ -1 & -1 \end{bmatrix}$	$\begin{bmatrix} +3 & 0 & 0 \\ -1 & +2 & 0 \\ -1 & -1 & +1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} +4 & 0 & 0 & 0 \\ -1 & +3 & 0 & 0 \\ -1 & -1 & +2 & 0 \\ -1 & -1 & -1 & +1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$

etc.

In each case, the number of vectors (columns, dummy variables) required is one less than the number of levels of the factor (= df for the factor). The rows correspond to the codes for each level in the experiment. The method of coding ensures orthogonality, as can be verified by computing cross products among the columns of the example matrices above. This pattern of coding produces correct sums of squares for given individual factors in an analysis-of-variance model. Other coding methods, involving particular contrasts, can be implemented, but the above methodology is easy to use and will provide valid tests for the total explained variation of particular factors for the present application.

Coding for multifactorial designs

We restrict detailed discussion to the example of a two-way crossed design. The linear model for the two-way crossed design is

$$y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \varepsilon_{ijk} \tag{C.8}$$

where y_{ijk} = the value of the kth replicate of the ith level of factor A and the jth level of factor B; μ = the parametric mean; A_i = the effect of the ith level of factor A; B_j = the effect of the jth level of factor B; $(AB)_{ij}$ = the interaction

effect of the ith level of factor A with the jth level of factor B; and ε_{ijk} = the random error associated with that replicate. For what follows, let $r = (a - 1)$ where a is the number of levels in factor A and $s = (b - 1)$ where b is the number of levels in factor B. The expectation of y_{ijk} under the full multiple-regression model can be written as

$$E(y) = \mu X_0 + t_{A1}X_{A1} + \ldots + t_{Ar}X_{Ar} + t_{B1}X_{B1} + \ldots + t_{Bs}X_{Bs} + t_{AB1}X_{AB1} + \ldots + t_{ABrs}X_{ABrs} \quad \text{(C.9)}$$

where the t's are individual regression coefficients corresponding to particular orthogonal vectors (X's), as shown in the one-way example above. The subscripts $A1, \ldots, Ar$ denote orthogonal vectors 1, . . . , r coding for factor A, subscripts $B1, \ldots, Bs$ denote orthogonal vectors $1, \ldots, s$ coding for factor B and subscripts $AB1, \ldots, ABrs$ denote orthogonal vectors 1, . . . , rs coding for the interaction between factors A and B.

As an example, consider a univariate two-factor orthogonal experiment in which there are three levels of factor A and two levels of factor B; n which is the number of replicates within each combination of levels of the two factors, is 3 in this case. The multiple-regression approach uses the following matrices:

$$\mathbf{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{123} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \\ y_{223} \\ y_{311} \\ y_{312} \\ y_{313} \\ y_{321} \\ y_{322} \\ y_{323} \end{bmatrix} \quad \mathbf{X} = \left[\begin{array}{ccc:c:cc} +2 & 0 & 1 & +2 & 0 \\ +2 & 0 & 1 & +2 & 0 \\ +2 & 0 & 1 & +2 & 0 \\ \hline +2 & 0 & -1 & -2 & 0 \\ +2 & 0 & -1 & -2 & 0 \\ +2 & 0 & -1 & -2 & 0 \\ \hline -1 & 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & 1 \\ \hline -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 \\ \hline -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 \\ \hline -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 \end{array}\right]$$

$$\mathbf{b} = \begin{bmatrix} t_{A1} \\ t_{A2} \\ \cdots\cdots \\ t_{B1} \\ \cdots\cdots \\ t_{AB11} \\ t_{AB21} \end{bmatrix}.$$

Note that the first column of matrix $\mathbf{X}$, X_0, has been omitted here. It consists of a column of 1's and corresponds to the centering of the data around the overall mean (i.e., for the estimation of μ in the multiple-regression model). This column is automatically added by the regression programs of statistical packages, unless the user specifies that the model must have an intercept of zero. So, in practice, this first column does not have to be written explicitly as one of the dummy variables.

The six different cells corresponding to the six combinations of levels of factor A and factor B are separated by horizontal solid lines. The vertical dotted lines in the $\mathbf{X}$ matrix separate the dummy variables coding for different terms in the model. The first two columns, X_{A1} and X_{A2}, are orthogonal vectors that code for factor A (three levels) and the third column, X_{B1}, codes for factor B (two levels). The last two columns code for the interaction term.

The columns of dummy values for interaction terms consist of the direct products of each of the orthogonal columns coding for one factor times each of the orthogonal columns coding for the other factor in the interaction (see Draper and Smith 1981). Before combining them into interaction dummy variables, it is essential that the variables used for any single main effect be linearly independent orthogonal vectors (Draper and Smith 1981).

This is a very efficient way to obtain dummy variables for the interaction term that are linearly independent of the variables coding for each of the main effects. The interaction term must not explain any portion of the variation in $\mathbf{Y}$ that is already explained by the main effects; it is on this basis that the interaction in an orthogonal (crossed) design is defined (e.g., Neter et al. 1996). In addition, by this method the number of orthogonal vectors required to code for the dummy variables corresponding to the interaction term is clearly equal to the number of degrees of freedom for that term in the model. Thus, in our example, $X_{AB11} = X_{A1}X_{B1}$ (column 4 in $\mathbf{X}$) and $X_{AB21} = X_{A2}X_{B1}$ (column 5 in $\mathbf{X}$). Similarly, for designs involving three or more factors, dummy variables coding for interaction terms can be obtained by this approach as the direct products of orthogonal dummy variables coding for each of the factors involved in the interaction.

As a final note, the above coding schemes for orthogonal vectors cannot be used with unbalanced data. Orthogonal vectors coding for least squares solutions using Type III or Type IV sums of squares would be appropriate for unbalanced designs (Searle 1987, Shaw and Mitchell-Olds 1993).