DOCUMENT RESUME

ED 271 493 | TM 860 402

AUTHOR	Robey, Randall R.; Barcikowski, Robert S.
TITLE	Type I Error in the Single Group Repeated Measures Design with Multiple Measures per Occasion.
PUB DATE	Apr 86
NOTE	51p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Analysis of Variance; Behavioral Science Research; Comparative Analysis; *Hypothesis Testing; Mathematical Models; *Monte Carlo Methods; *Multivariate Analysis; Observation; Predictor Variables; *Research Design; Statistical Analysis; *Statistical Significance; Statistical Studies
IDENTIFIERS	*Repeated Measures Design; Sphericity Tests; *Type I Errors

ABSTRACT

This paper reports the results of a Monte Carlo investigation of Type I errors in the single group repeated measures design where multiple measures are collected from each observational unit at each measurement occasion. The Type I error of three multivariate tests were examined. These were the doubly multivariate F test, the multivariate mixed model F test, and a recently proposed multivariate extension of Mauchly's sphericity test. The study examined the efficacy of using the latter test as preliminary test to evaluate the tenability of the multivariate sphericity assumption made by the multivariate mixed model F. The study also examined the impact of violations of the multivariate sphericity assumption upon the Type I error control of the multivariate mixed model F. The results indicate an analysis strategy for the researcher confronted with the task of analyzing data gathered using this research design. Example data sets are presented to illustrate three likely patterns of behavioral data. The data analysis strategy is discussed as it relates to each data set. (Author)

Type I Error in the Single Group Repeated Measures Design with Multiple Measures per Occasion

Randall R. Robey

Southern Illinois University

Robert S. Barcikowski

Ohio University

A paper presented at the annual meeting of the American Educational Research Association, San Fransico, April, 1986.

Abstract

This paper reports the results of a Monte Carlo investigation of Type I errors in the single group repeated measures design where multiple measures are collected from each observational unit at each measurement occasion. The Type I error of three multivariate tests were examined. These were the doubly multivariate F test, the multivariate mixed model F test, and a recently proposed multivariate extension of Mauchly's sphericity test. The study examined the efficacy of using the latter test as preliminary test to evaluate the tenability of the multivariate sphericity assumption made by the multivariate mixed model F. The study also examined the impact of violations of the multivariate sphericity assumption upon the Type I error control of the multivariate mixed model F.

The results indicate an analysis strategy for the researcher confronted with the task of analyzing data gathered using this research design. Example data sets are presented to illustrate three likely patterns of behavioral data. The data analysis strategy is discussed as it relates to each data set.

Type I Error in the Single Group Repeated Measures Design with Multiple Measures per Occasion

Introduction

Behavioral scientists often find it necessary, or desirable, to measure a dependent variable on some observational unit (usually subjects) under more than one treatment condition. A research design structured such that a sample of units is observed on more than one occasion has been described by a variety of terms: treatment by subjects (Lindquist, 1953), two-way mixed model (Scheffé, 1959), single factor repeated measures (Winer, 1971), design on the occasions (Bock, 1975), randomized block (Kirk, 1982), and single factor within-subjects (Keppel, 1982).

The necessity of using repeated measures designs often occurs since much data naturally exist in this form (McHugh, Sivanich and Geisser, 1961; Rich, 1983). This occurs since many behavioral and physiological attributes are of primary interest as they vary over time, and under different treatment conditions.

Behavioral scientists may find a repeated measures design desirable since subjects serve as their own control. This has the effect of reducing that part of error variance which can be attributed to between group subject heterogeneity in a factorial design. As a result, statistical power is increased relative to the factorial alternative and fewer subjects are required for study.

Data gathered in a single group repeated measures design can be inferentially tested using two distinctly different methods of analysis. These two analyses can yield different results depending upon the nature of the data. The first analysis is the classic mixed model repeated measures analysis of variance (Scheffé, 1959). The second analysis tests the same hypothesis, but is conducted as a multivariate analysis of variance (Timm, 1975).

Extensions of both the mixed model test and the multivariate test have been proposed for repeated measures research designs in which more than one dependent variable is observed in each subject on each occasion. The extension of the mixed model test is termed the 'multivariate mixed model' test (Bock, 1975). In this analysis, the usual mixed model sums of squares are replaced with sums of squares and cross products matrices which contain information on all of the dependent variables. In addition to the assumptions of the mixed model test, the multivariate mixed model test rests upon the even more stringent assumption of sphericity in each of the dependent variables and constant variance-covariance structure among the dependent variables across occasions, i.e., multivariate sphericity, (Timm, 1980).

The extension of the multivariate test is termed the 'doubly multivariate' test (Bock, 1975). This term refers to the fact that the multivariate approach to repeated measures is used to analyze multiple dependent measures on each unit at each occasion. In addition to the assumptions for the multivariate test, the doubly multivariate test assumes constant structure of the variance-covariance matrices for the dependent variables across occasions. In this analysis, n must exceed the hypothesis degrees of freedom (i.e., (k-1)p, where k is the number of measurement occasions, and p is the number of dependent variables).

Both tests have the advantage of considering all of the dependent variables, and their inter-relationships, at the same time. This means that a researcher can examine how the dependent variables are affected, as a whole, by the treatments. If the dependent variables accurately represent the theoretical construct, a researcher is in an enhanced position to discuss that construct relative to a colleague testing a single dependent variable, or testing multiple dependent variables independently.

Statistical Models: One Dependent Variable per Occasion

The Multivariate Model

Using the full rank model, the general form of the multivariate null hypothesis for a design with g groups, k repeated measures, and one dependent variable per occasion is:

$$H_0: \; ABC' = D$$

where A is a g-1 x g contrast matrix representing the between group hypothesis, B is a g x k matrix of cell means, and C is a k-1 x k contrast matrix representing the within factor hypothesis (Timm, 1975; Timm and Carlson, 1973). This paper is concerned with the single group repeated measures design where there is no between group hypothesis. Since, in this case g = 1, the matrix A is a scalar set at unity. The matrix D is usually a null matrix.

The multivariate sums of squares and cross products matrices (SSCP) for the hypothesis (H) and error (E) are given by:

$$H = CB'(X'X)^{-1}BC' \quad \text{, and} \qquad (1)$$

$$E = C\,(Y-\hat{Y})'(Y-\hat{Y})\,C' \qquad , \qquad (2)$$

where B is a 1 x k vector of the treatment means over the k occasions, and X is a design matrix. Here, X is an n x 1 vector of ones, where n is the number of subjects. The matrix C is the k-1 x k contrast matrix, and Y is an n x k matrix containing the n vectors of observations. The matrix $\hat{Y}$ is obtained using $\hat{Y}$ = XB.

The omnibus multivariate repeated measures hypothesis can be tested using

$$\Lambda = |E| \;/\; |E + H| \qquad (3)$$

where Λ is Wilks's likelihood ratio criterion (Wilks, 1932) with k-1, 1, and n-k+1 degrees of freedom, and |E| denotes the determinant of the matrix E. In the single group case, other multivariate tests (i.e., Roy's largest root, etc.) all share this distribution.

A multivariate F statistic can be obtained by

$$F = [(1-\Lambda)/\,\Lambda\,]\;(v_1/v_2) \qquad (4)$$

where v_1 are the k-1 hypothesis degrees of freedom, and v_2 are the n-k+1 error degrees of freedom. When k = 2, the mixed model test and the multivariate test are identical.

<u>The Mixed Model</u>

If the contrasts in C are row-wise orthonormal, the omnibus mixed model test statistic can be obtained from Equations 1 and 2. The omnibus mixed model test statistic obtained through the multivariate model is given by

$$F = [tr(H)/v_1] \;/\; [tr(E)/v_2] \qquad , \qquad (5)$$

where tr is the trace of a matrix, v_1 are the k-1 hypothesis degrees of freedom, and v_2 are the (k-1)(n-1) error degrees of freedom.

Alternatively, the mixed model SSCP matrices for the hypothesis (T) and error (M) are given by

$$T = [(C\#1)Y]'\;[(C\#1)(C\#1)']^{-1}\;[(C\#1)Y]\;,\ \text{and} \qquad (6)$$

$$M = [(C\#Z)Y]'\;[(C\#Z)(C\#Z)']^{-1}\;[(C\#Z)Y]\;. \qquad (7)$$

The matrix 1 is a vector of n ones. The matrix C is a k-1 x k matrix of simple contrasts on the repeated measures. The matrix Z is an n-1 x n contrast matrix representing the main effect hypothesis on the subjects as a random effect. The operation # denotes the Kronecker product of two matrices. Note that the Kronecker product of C and 1 defines the treatment effect, and that the Kronecker product of C and Z defines the treatment by subjects interaction effect. The nk x 1 vector Y contains the dependent variable observations ordered by treatments, and subjects ordered within the treatments. The design matrix X is actually an identity matrix of order n. These alternate expressions are useful in developing the $p > 1$ generalization of the mixed model.

Both T and M in Equations 6 and 7 are scalars. These are the traditional mixed model sums of squares for hypothesis and error, respectively. They are divided by their appropriate degrees of freedom [k-1 and (n-1)(k-1), respectively] to obtain the mean square values for the traditional ratio of mean squares. The F test given by this ratio of mean squares is equivalent to Equation 5.

The Mixed Model and Sphericity. E.senhart (1947) stated that the mixed model F test was valid only when the correlations among the repeated measures were all zero. This assumption was later modified to state that the variances among the repeated measures must be equal, and that the covariances among the repeated measures must be equal. Said another way, the test assumes zero, or constant, correlations among the repeated measures. This latter condition has been termed 'uniformity' (Davidson, 1972), as well as 'compound symmetry' (Winer, 1971, p. 136).

Huynh and Feldt (1970) and Rouanet and Lépine (1970) showed that uniformity was a sufficient, but not a necessary condition for an exact distribution of F in Equation 5. They demonstrated that a less restrictive structure of the variance-covariance matrix is both necessary and sufficient for the test in Equation 5 to be distributed as F. Huynh and Feldt (1970) referred to this condition as 'sphericity', while Rouanet and Lépine (1970) used the term 'circularity'. In this paper the term 'sphericity' is used. Both Huynh and Feldt (1970) and Rouanet and Lépine (1970) showed that uniformity always implies sphericity, but that sphericity does not always imply uniformity.

Sphericity is held if, and only if,: (1.) the distribution of the repeated measures observations in the matrix Y is multivariate normal, and (2.) $C\Sigma C' = s^2 I$, where C is a k-1 x k orthonormal contrast matrix, Σ is the population variance-covariance matrix, and I is an identity matrix of order k-1. The element s^2 is a scalar > 0 which represents the common population variance on each of the contrasts in C (Boik, 1981).

Testing sphericity. Mauchly (1940) proposed a test to determine if $C\Sigma C' = s^2I$ is a tenable assumption for a sample of data. Mauchly's test for uniformity addresses the hypothesis that a sample of variates was selected from a population where the variances are all equal, and the covariances are all zero (i.e., s^2I). Therefore, Mauchly's test is applied to the CSC' matrix. Note that this special case of compound symmetry (i.e., zero correlations) for the CSC' matrix represents, by definition, sphericity among the original variables. Mauchly's test is defined by Huynh and Feldt (1970) as

$$X^2 = -(n-1)\ln(w)d \qquad (8)$$

$$d = 1 - (2(k-1)^2 + (k-1) + 2) / [6(k-1)(n-1)] \qquad (9)$$

$$w = |CSC'| / [tr(CSC') / (k-1)]^{-1} . \qquad (10)$$

The test statistic in Equation 8 is evaluated as Chi-square with [(k-1)k/2] -1 degrees of freedom. Nargarsenker and Pillai (1973), among others, have described the exact distribution of the criterion w, thus eliminating the need for using the Chi-square approximation.

Mauchly's (1940) test is applied to a sample of data before conducting the repeated measures analysis of variance. It follows that rejection of the null hypothesis indicates that a mixed model test would not be valid (Rogan, Keselman, and Mendoza, 1979.) Therefore, the data require an alternate test statistic. Three studies have investigated the efficacy of Mauchly's test as a preliminary test of sphericity.

Huynh and Mandeville (1979) conducted a Monte Carlo investigation of the Type I error performance of Mauchly's test under conditions of normality and non-normality. The non-normal conditions were variations of kurtosis (i.e., light tailed and heavy tailed). They found that under normality, Mauchly's test showed good agreement between the actual and nominal alpha levels (i.e., the test was neither conservative nor liberal). Under light tailed distributions (i.e., leptokurtosis), the test performed conservatively with discrepancy of the actual and nominal alpha levels increasing as n increased. The test performance under heavy tailed distributions (i.e., platykurtosis) was excessively conservative. Huynh and Mandeville (1979) concluded that Mauchly's test was useful as a preliminary test of sphericity for normal and light tailed distributions. The authors stated that since behavioral data will likely vary from normality in the light tailed direction, Mauchly's test is a useful preliminary test of sphericity for behavioral data.

This issue is clouded by a pair of studies which concluded that Mauchly's test does not need to be conducted. Rogan, Keselman and Mendoza (1979) and Keselman, Rogan, Mendoza and Breen (1980) reported results which appear to represent different aspects of the same study. In these two studies, the authors varied not only normality, but also departures from sphericity, in a Monte Carlo study of the Type I error rate and the statistical power of Mauchly's test. One of the conditions of non-sphericity was designed to represent what would probably be a trivial departure from sphericity for the practitioner. Another non-sphericity condition represented what Huynh and Feldt (1976) have suggested is generally the lower limit of non-sphericity in educational data. Two additional conditions represented more severe violations of the sphericity assumption.

Rogan et al. (1979) and Keselman et al. (1980) found that Mauchly's test was very sensitive to departures from normality, and to departures from the null hypothesis: H_0: $C\Sigma C' = s^2 I$. They suggested that the rejection rate for Mauchly's test of sphericity was so high, even for mild departures, that a researcher might as well assume non-sphericity when analyzing repeated measures data.

Thomas (1983) commented that the excessive conservatism found by Keselman et al. (1980) might be attributed to the relatively large n (i.e., 39) used in that study. This observation is supported by Huynh and Mandeville's (1979) finding that, under non-normality, the actual alpha level decreased as n increased. This notion led Thomas to suggest that preliminary testing of sphericity may be a useful analytical strategy when n is relatively small.

<u>The sphericity index</u>. Under a violation of sphericity, F is not distributed with (k-1) and (k-1)(n-1) degrees of freedom. Box (1954) demonstrated that a multiplicative correction factor for these degrees of freedom is defined by

$$\epsilon = \frac{\left(\sum_{i=1}^{K-1} \lambda_i\right)^2}{(k-1)\sum_{i=1}^{K-1} \lambda_i^2} \qquad (1i)$$

where the λ_i (i = 1, 2,...k-1) are the k-1 eigenvalues of $C\Sigma C'$.

That is, Box (1954) showed that the mixed model F is approximately distributed as $F[(k-1)\epsilon, (n-1)(k-1)\epsilon]$. Note that if the λ's are all equal, ϵ equals 1; and equals a fraction otherwise. Rogan, Keselman, and Mendoza (1979) gave an equivalent form of Equation 11 as

$$\epsilon = \frac{[tr(C\Sigma C')]^2}{(k-1)\ tr(C\Sigma C')^2} \quad . \qquad (12)$$

The value of ϵ is an index of sphericity. Geisser and Greenhouse (1958) showed that ϵ is bounded by 1 and 1/(k-1). Sphericity is held when $\epsilon = 1$, and is maximally violated when $\epsilon = 1/(k-1)$. Imhof (1962) recognized that if $C \Sigma C'$ is positive definite, ϵ can approach, but not equal, 1/(k-1).

Statistical Models: Multiple Dependent Variables per Occasion

The Doubly Multivariate Model

The doubly multivariate repeated measures analysis of variance is a direct extension of the multivariate approach to repeated measures analysis of variance for the single dependent variable case. The dimensions of the matrix Y are expanded to n x kp, where p is the number of qualitatively different dependent variables. This simply means that the n x k matrices of observations for each of the dependent variables are concatenated horizontally to form Y. The doubly multivariate sums of squares and cross products matrices are given by

$$H^* = (I\#C)B'\ (X'X)^{-1}\ B(I\#C)' \quad \text{and,} \qquad (13)$$

$$E^* = (I\#C)\ (Y-\hat{Y})'(Y-\hat{Y})\ (I\#C)' \quad . \qquad (14)$$

where I is a p x p identity matrix. The I#C operation is a convenient method for developing a contrast matrix for the doubly multivariate hypothesis. As a result of this operation, the same contrasts will be applied to each of the dependent variables. However, it should be noted that a contrast matrix representing different hypotheses for different dependent variables can be constructed.

Once again, the multivariate test statistic Λ, is found using $\Lambda = |E^*| \ / \ |E^* + H^*|$. A multivariate F statistic can be found using Equation 4 with (k-1)p hypothesis degrees of freedom, and M + 1 - ((k-1)p/2) error degrees of freedom, where M = (n-1) - ((k-1)p/2).

The Multivariate Mixed Model

When k = 2, the multivariate mixed model and the doubly multivariate F statistics are the same. When $k \geq 3$, the SSCP matrices for the mixed model multivariate test cannot be derived directly from the doubly multivariate H* and E* matrices by using a convenient matrix operation such as the trace of a matrix. However, Timm (1980, p. 73) has shown that the elements of the mixed model multivariate mean square matrices can be obtained by hand from the doubly multivariate mean square matrices. As in the single dependent variable analysis, this derivation can be made only when the contrasts in C are orthonormal. Timm's procedure can also be applied to SSCP matrices. In essence, one sums the contrast variances for each dependent variable, and sums the covariances on each contrast for each pair of dependent variables.

In the p > 1 generalization of Equations 6 and 7, the multivariate mixed model SSCP matrices for the hypothesis (T*) and error (M*) are obtained using

$$T^* = [(C\#1)Y]' \, [(C\#1)(C\#1)']^{-1} \, [(C\#1)Y] \quad \text{, and} \qquad (15)$$

$$M^* = [(C\#Z)Y]' \, [(C\#Z)(C\#Z)']^{-1} \, [(C\#Z)Y] \; . \qquad (16)$$

Here, the matrix 1 is a row vector of n ones, and C is a k-1 x k contrast matrix. The matrix of observations for the multivariate mixed model is arranged in the mixed model with the additional p row vectors of observations being horizontally concatenated.

The sums of squares and cross product matrices in Equations 15 and 16 are used in $|M^*| \,/\, |M^* + T^*|$ to obtain the multivariate test statistic Λ. A multivariate F statistic is then given by

$$F = [(1 - \Lambda^{s^{-1}}) / \Lambda^{s^{-1}}] \,/\, (v_1/v_2) \qquad (17)$$

where

$$s = \sqrt{\frac{p\,(k-1) - 4}{p^2 + (k-1)^2 - 5}} \qquad (18)$$

when (k-1)p > 2, and equals unity otherwise. In Equation 17,

v_1 and v_2 are used to denote the mixed model multivariate hypothesis and error degrees of freedom respectively. The F statistic has (k-1)p hypothesis degrees of freedom, and

$$(n-1)(k-1)s + 1 - (k-1)p/2 \qquad (19)$$

error degrees of freedom.

Multivariate Sphericity. The multivariate mixed model analysis assumes not only sphericity of the variance-covariance matrix associated with each of the dependent variables, but that the variance-covariance structure among the dependent variables, across occasions, is the same (Timm, 1980). That is,

$$V_1 = V_2 = V_3 \ldots = V_i = V,$$

where V_i (i = 1,2,...k) is the variance-covariance matrix for the dependent variables at the ith occasion, and where V is the common variance-covariance matrix of the dependent variables for all k occasions.

Testing multivariate sphericity. Thomas (1983) extended Mauchly's test of sphericity for p = 1 to the any p case. Let D = (1/p)M*. Further, let t_i be the natural log of the ith eigenvalue in Equation 14, and let u_i be the natural log of the ith eigenvalue of D. Then Thomas's extension of Mauchly's test is given by

$$\chi^2 = ng[(k-1)\sum_{i=1}^{K-1} u_i - \sum_{i=1}^{K-1} t_i], \qquad (20)$$

where g is the number of groups. This Chi-square is evaluated with

$$\frac{p(k-2)\ [p(k-1) + p + 1]}{2} \qquad (21)$$

degrees of freedom.

Comparison of the Two Models

As in the single dependent variable case, when k = 2, the multivariate mixed model and the doubly multivariate analyses yield identical test statistics. However, when $k \geq 3$, the results of the two analyses can vary depending upon the nature of the data being tested.

Jensen (1982) conducted an empirical analysis of the statistical power of the multivariate mixed model test, and of the doubly multivariate test. Jensen was able to define asymptotic bounds for the uniform power dominance of one, or the other, test. However, Jensen's results apply only to the restrictive condition of independence among the p measures (i.e., zero correlations). In order to discuss Jensen's findings, let R be a correlation matrix of the p measures over the k occasions, and let C be a k-1 x k orthonormal contrast matrix. Further, let Q = (I#C)(R-I)(I#C)', where I is an identity matrix of order p. Jensen (1982) was able to show that under independence of the measures, when Q is positive definite, the doubly multivariate test will uniformly dominate the multivariate mixed model test, in terms of statistical power. This power domination is reversed when Q proves to be negative definite. For those cases where Q is indefinite, Jensen (1982) provided a method for calculating the estimated bounds of the two tests.

Conclusions

Although the mixed model test rests upon the stringent sphericity asssumption, several lines of investigation have provided the data analyst with some analysis options. After considering Type I and Type II error, Barcikowski and Robey (1984a, 1984b) recommended the application of both the multivariate test and the mixed model test (corrected for ϵ). They suggest splitting the a priori alpha equally between the tests.

However, less information is available concerning the p > 1 case, and a preferred analysis strategy is not clearly indicated. Based upon literature concerning sphericity for p = 1 (i.e., Huynh and Mandeville, 1979; and Keselman, et al., 1980), some example problems, and a small simulation study, Thomas (1983) came to the following conclusions. The choice between the doubly multivariate model and the multivariate mixed model analyses for a test statistic rests upon the outcome of the extended Mauchly test (i.e., the Mauchly-Thomas test). If the Mauchly-Thomas Chi-square is significant, the practitioner should submit the data to the doubly multivariate analysis since the multivariate sphericity assumption of the multivariate mixed model is not satisfied. If the extended Mauchly test turns out to be insignificant, the multivariate mixed model should be employed since it seems more powerful than the doubly multivariate analysis considering the results of the simulation study. However, Thomas (1983) cautioned that this apparent power differential and the performance characteristics of the

Mauchly-Thomas test require further study.

Purposes of the Present Study

The purpose of the present Monte Carlo investigation was two fold. First, the study was designed to estimate the Type I error rate of the Mauchly-Thomas Chi-square statistic. Considering the results of Rogan et al. (1979) and of Keselman et al. (1980) for Mauchly's original test in the $p = 1$ case, this aspect of the study was extended to include an estimate of the sensitivity of the Mauchly-Thomas test to departures from its null hypothesis.

Secondly, the present investigation sought to estimate the Type I error rate of the multivariate mixed model F test since its sampling distribution is not known under departures from the multivariate sphericity assumption. Given Huynh's (1978) finding that the mixed model test for the $p = 1$ design is not robust with respect to even small departures from sphericity, it seems reasonable to examine the multivariate extension of this test for the $p > 1$ design. Since the sampling distribution of the doubly multivariate F test is known, and since it does not assume multivariate sphericity, this statistic was calculated in all simulations as one accuracy check of the simulation procedure.

Method

Independent Variables

A Monte Carlo analysis was conducted for each combination of the various levels of the following independent variables. Each of these combinations are referred to as a Monte Carlo problem. The three test statistics of interest (i.e., the multivariate mixed model F, the doubly multivariate F, and the Mauchly-Thomas Chi-square) were calculated from 2000 samples for each Monte Carlo problem.

Occasions. The number of the k occasions comprising the research design was varied at 3, 5, and 7.

Dependent variables. The number of dependent variables in the design was varied at 2 and 3.

Violations of multivariate sphericity. Halperin (1976) advocated improving the efficiency of Monte Carlo studies through what he called 'variance reduction'. In repeated measures designs at the $p = 1$ level, such variance reduction is accomplished by generating data from a $C\Sigma C'$ matrix which has been reduced to its canonical form (i.e., a diagonal matrix of k-1 eigenvalues.) This practice is much more efficient than generating data from the original variance-covariance matrix for the k occasions. However, the writers are not aware of an analogous canonical form for a $C\Sigma C'$ matrix which is actually a super matrix. Therefore, the data were generated from the kp x kp variance-covariance matrix of the original variables. The structure of the variance-covariance matrices was fashioned after Timm (1980). That is, a variance-covariance matrix was organized by k across the levels of p.

The source of the violations of multivariate sphericity was limited to the k x k component variance-covariance matrix for each dependent variable along the diagonal. In the first condition, multivariate sphericity was held, i.e., $\epsilon = 1$ for all of the dependent variables. In the second condition, each dependent variable demonstrated a 'mild' departure from sphericity, i.e., $\epsilon = .95$. This condition represented small departures from multivariate sphericity which many researchers would probably ignore. In the third condition, ϵ for each of the dependent variables equaled .75. This condition represented a'moderate' departure from multivariate sphericity. The selection of this value is based upon Huynh and Feldt's (1976) finding that the lower limit of ϵ in practice is approximately .75. Copies of the matrices used in this study are available upon request.

Correlations among the dependent variables across occasions. The constant correlations among the dependent variables, across occasions, were varied at r = .2, .5, and .8. All other correlations among the various levels of k and p were set at 0. The correlations were represented in the diagonal elements of the k x k component matrices of each super matrix.

Sample sizes. Based upon Davidson (1972), the number of observations in the research design was varied at p(k-1) + 3, p(k-1) + 10, and p(k-1) + 20. These sample sizes represented small, moderate, and moderately large n's relative to p(k-1).

Nominal alpha levels. The nominal alpha level for all tests was varied at .01 and .05.

Dependent Variables

The Monte Carlo dependent variable was the proportion of the 2000 calculated test statistics which exceed the tabled critical value for that test. For each Monte Carlo problem, a proportion of exceedance was calculated for each of the three test statistics being evaluated.

A check on the validity of the Monte Carlo analysis was conducted for those Monte Carlo problems where multivariate sphericity was held. This check, given by Olson (1973), is called the Monte Carlo critical value. The definition of the Monte Carlo critical value may best be explained by way of example. For instance, when alpha (α) is set at .05 and 1000 samples are observed, the Monte Carlo critical value is the mean of the 50th and 51st largest of the 1000 calculated test statistics. This value should approximate the tabled critical value for alpha at .05. A general definition of the Monte Carlo critical value states that it is the mean of the ($N\alpha$)th and the ($N\alpha$ +1)th largest of the N calculated test statistics.

Data Generation

A FORTRAN subroutine, GGNSM, from the International Mathematical and Statistical Libraries, Inc. (IMSL, 1982) was used to generate multivariate normal data for each of the variance-covariance matrices. GGNSM first generates multivariate normal vectors of random numbers, N(0,1). Then using Cholesky decomposition, the input variance-covariance matrix is decomposed to an upper triangular matrix, U, such that $UU' = \Sigma$. The N(0,1) vectors of data in some matrix, say Z, are then transformed to $N(0,\Sigma)$ through ZU'.

The data for a given variance-covariance matrix were generated in one execution of a program using multiple calls to GGNSM. The initial seeds for GGNSM were selected from a table of random numbers. GGNSM generates new seeds upon successive calls. The output data were written directly to tape.

Calculation of the Statistical Tests

The three statistical tests of interest were calculated via a G level FORTRAN program developed specifically for this purpose. Photocopies of the program are available upon request.

Statistical Hypotheses

Null Hypotheses. The null hypothesis for those Monte Carlo problems where the independent variable, nominal alpha, was set at .01, is written as

$$H_0: \quad P = .01,$$

where P represents a proportion. Likewise, the null hypothesis for those Monte Carlo problems where the independent variable, nominal alpha, was set at .05, is written as

$$H_0: \quad P = .05,$$

Alternate Hypotheses. The alternate hypothesis for those Monte Carlo problems where the independent variable, nominal alpha, was set at .01 is written as

$$H_A: \quad P \neq .01,$$

Similarly, the alternate hypothesis for those Monte Carlo problems where the independent variable, nominal alpha, was set at .05 is written as

$$H_A: \quad P \neq .05,$$

Statistical Analyses

It was decided that departures of $\pm$ .01 from the nominal alpha of .01, and that departures of $\pm$ .02. from the nominal alpha of .05 would constitute meaningful differences between nominal and actual alpha levels.

A two-tailed test for proportions described by Cohen (1977, p. 213) was used to analyze the results of the Monte Carlo problems. In those Monte Carlo problems where the null hypothesis of multivariate sphericity was known to be false, the proportions test was not conducted for the Mauchly-Thomas Chi-square results. Indeed, for these research problems, the number of times which the null hypothesis was rejected provided some idea of the statistical power of the Mauchly-Thomas test. The a priori alpha level for all applications of Cohen's (1977) test was set at .01.

The desired minimal statistical power for all applications of the proportions test was set at .80. Following the method for establishing sample size described by Cohen (1977), it was determined that 1678 observations were needed for H_0: P = .01, and that 1635 observations were needed for H_0: P = .05. As a matter of convenience, 2000 observations were collected for each Monte Carlo problem. As a result, statistical power exceeded .80 in all of the analyses.

Results

This section is organized in two major segments. The actual alpha rates for the various Monte Carlo problems are presented in the first segment. These results are summarized in Tables 2 through 10. In the second segment, the Monte Carlo critical value results are reported. These results are summarized in Tables 11 through 16. All tables are found in the Appendix.

Table 1 can be used to facilitate the interpretation of

Tables 2 to 10. Table 1 contains the calculated sample sizes used for each combination of k (3, 5 and 7), p (2 and 3), and n [(k-1)p+3, (k-1)p+10, and (k-1)p+20].

Actual Alpha Rates

The doubly multivariate test. The actual alpha levels for the doubly multivariate F tests are found in Tables 2 through 4. The actual alpha levels in Table 2 were obtained in Monte Carlo problems where multivariate sphericity was maintained. The actual alpha levels in Tables 3 and 4 were obtained in Monte Carlo problems which, respectively, represented the conditions of mild and moderate departures from multivariate sphericity.

As expected, the doubly multivariate F test controlled Type I error across the range Monte Carlo problems. However, on three isolated occurrences, the two-tailed proportions test rejected its null hypothesis. The first occurrence is found in Table 3 with the remaining two occurrences appearing in Table 4. The three Monte Carlo problems are defined as follows.

	Departure from Multivariate Sphericity	Correlation Among p Across k	k	p	n	Nominal Alpha
1.	mild	.2	5	2	(k-1)p+20	.01
2.	moderate	.2	3	3	(k-1)p+20	.05
3.	moderate	.8	5	3	(k-1)p+20	.01

The only characteristic that all three Monte Carlo problems share is sample size, i.e., n = (k-1)p+20. Otherwise, these three problems do not seem to be linked by an overall pattern, nor do they seem to represent a pattern among the independent variables. Although the absolute value of the difference between the actual alpha level and the nominal alpha level are significantly different in all three problems, each of the actual alpha levels are within the appropriate specified tolerances, i.e., between .00 and .02 for nominal alpha at .01, or between .03 and .07 for nominal alpha at .05. The three rejections in a total of 324 tests represent a Type I error rate for the two-tailed proportions test of .00925. This is in good agreement with the a priori alpha level set at .01. It seems reasonable to conclude that the doubly multivariate F test demonstrated appropriate Type I error control.

The multivariate mixed model test. The actual alpha levels for the multivariate mixed model F tests are found in Tables 5 through 7. The actual alpha levels in Table 5 were obtained in Monte Carlo problems where multivariate sphericity was maintained. The actual alpha levels in Tables 6 and 7 were obtained in Monte Carlo problems which, respectively, represented the conditions of mild and moderate departures from multivariate sphericity.

As can be seen in Table 5, the multivariate mixed model F test controlled Type I error when multivariate sphericity was maintained. The null hypothesis of the two-tailed proportions test was rejected on one occasion. The Monte Carlo problem in this case was defined by k at 3, p at 3, the correlation among the dependent variables across occasions at .5, n at (k-1)p+3, and nominal alpha at .01. The actual alpha of .004 for this Monte Carlo problem is within the tolerance specified for a nominal alpha of .01, i.e., between .00 and .02 for nominal alpha at .01. This Monte Carlo problem does not seem to characterize any pattern among the independent variables. This single rejection in a total of 108 Monte Carlo problems represents a Type I error rate of .00925 for the two-tailed proportions test. This is in good agreement with the a priori alpha level of .01.

The actual alpha levels obtained for the multivariate mixed model F test under the condition of a mild departure from multivariate sphericity demonstrated good control of Type I error. Only one Monte Carlo problem resulted in a rejection of the null hypothesis of the two-tailed proportion test. The Monte Carlo problem in this case was defined by k at 3, p at 3, the correlation among the dependent variables across occasions at .5, n at (k-1)p+10, and nominal alpha at 05. The actual alpha level of .071 for this Monte Carlo problem represents a modest departure from the specified tolerance for nominal alpha set at .05, i.e., between .03 and .07 for nominal alpha at .05. In general, when compared to the results in Tables 2 through 5, a fewer number of the actual alpha levels in Table 6 are less than the nominal alpha level. That is, under a mild departure from multivariate sphericity, the actual alpha levels for the multivariate mixed model F test were, in general, slightly elevated for each combination of the independent variables.

A large majority of the actual alpha levels reported in Table 7 represent a significant difference from the corresponding nominal alpha levels. The exceptions, those 14 Monte Carlo problems where Type I error was controlled, are summarized in Table 17. The 14 Monte Carlo problems do not seem to represent any pattern among the 108 problems which come under the condition of a moderate departure from multivariate sphericity.

Interestingly, the actual alpha level for each of these 108 Monte Carlo problems exceeds the nominal alpha. Further, it should be noted that while most of the actual alpha levels significantly depart from their nominal alpha level, none of the departures are drastic in nature. In Table 7, the largest actual alpha for a nominal alpha of .01 was .033, and the largest actual alpha for a nominal alpha of .05 was .086.

The Mauchly-Thomas chi-square test. The actual alpha levels for the Mauchly-Thomas Chi-square tests are found in Tables 8 through 10. The actual alpha levels in Table 8 were obtained in Monte Carlo problems where multivariate sphericity was maintained. The proportions of rejection in Tables 9 and 10 were obtained in Monte Carlo problems which, respectively, represented the conditions of mild and moderate departures from multivariate sphericity. Because the null hypothesis of multivariate sphericity was known to be false, the two-tailed proportions test was not conducted on the results found in Tables 9 and 10. Rather, the proportions of rejection in these two tables provide an estimate of the statistical power of the Mauchly-Thomas Chi-square statistic.

The actual alpha levels of the Mauchly-Thomas Chi-square statistic, under the null hypothesis of multivariate sphericity, are found in Table 8. Every actual alpha in this table is significantly greater than its nominal alpha. Two patterns in Table 8 are clear. First, the Type I error rate decreases as the sample size becomes larger. Secondly, the Type I error rate increases as k and/or p increases.

The proportions of rejection for the Mauchly-Thomas Chi-square statistic under a mild departure from multivariate sphericity, found in Table 9, demonstrate similar patterns. That is, the proportions of rejection decrease with an increase in sample size, and increase with an increase in k and/or p. However, as sample size becomes larger, the drop in the proportions of rejection is substantially less. This suggests that under a mild departure from the null hypothesis, the Mauchly-Thomas Chi-square statistic behaved much like it did under the null hypothesis with an attenuation of the sample size effect.

Table 10 contains the actual alpha levels of the Mauchly-Thomas Chi-square statistic under a moderate departure from multivariate sphericity. Once again, it can be seen in Table 10 that actual alpha increased as k and/or p increased. However, with few exceptions, as sample size increased, the actual alphas increased as well. This reflects the usual relationship between effect size (i.e., the degree of departure from multivariate sphericity) and sample size. That is, larger n's are better able to detect departures from the null hypothesis. As can be seen, with an n of (k-1)p+20, the Mauchly-Thomas Chi-square statistic is very sensitive to this moderate departure from multivariate sphericity, i.e., power is quite good. Even for smaller sample sizes, the power of the Mauchly-Thomas Chi-square statistic seems to be good for the higher combinations of k and p.

Monte Carlo Critical Values

Monte Carlo critical values were calculated for all three multivariate tests in Monte Carlo problems where multivariate sphericity was maintained. This condition was selected for the calculation of Monte Carlo critical values because it is the only condition where the assumptions of all three multivariate tests are satisfied, and where their null hypotheses are all true. The Monte Carlo critical values are found in Tables 11 through 16.

The doubly multivariate test. The Monte Carlo critical values, and the tabled critical values, for the doubly multivariate F statistic are found in Tables 11 and 12. Table 11 contains the critical values for those Monte Carlo problems where p was set at 2, and Table 12 contains the critical values for those Monte Carlo problems where p was set at 3.

In general, the Monte Carlo critical values for the doubly multivariate F statistic reasonably approximate their tabled critical value counterparts. The agreement is generally closer for those values where nominal alpha equals .05. This is because the critical values for nominal alpha at .01 are located further out in the tail of the sampling distribution. The largest discrepancies between a Monte Carlo critical value and a tabled critical value are associated with the smallest sample size. In each of these cases, the F distribution is based on only three error degrees of freedom. The sharp slope and the sparseness of the tail in such a sampling distribution explains these larger differences. The median percentage of discrepancy for the doubly multivariate test with nominal alpha set at .01 was 3.77 between a minimum of 0 and a maximum of 39.7. The median percentage of discrepancy for the doubly multivariate test with nominal alpha set at .05 was 44 between a minimum of 0 and a maximum of 16.01. In comparing Olson's (1973) Monte Carlo results to those of the present study, one should keep in mind that this study does not contain a condition where nominal alpha equals .10, and that one sample size is much smaller than those used by Olson. Considering these factors, these results compare favorably with those of Olson's (1973) Monte Carlo study.

The multivariate mixed model test. The Monte Carlo critical values, and the tabled critical values, for the multivariate mixed model F statistic are located in Tables 13 and 14. Table 13 contains the critical values for those Monte Carlo problems where p was set at 2, and Table 14 contains the critical values for those Monte Carlo problems where p was set at 3.

In general, the Monte Carlo critical values for the multivariate mixed model F statistic reasonably approximate their tabled critical value counterparts. Once again, the agreement is generally better for those Monte Carlo problems where nominal alpha was set at .05. The median percentage of discrepancy for the multivariate mixed model test with nominal alpha set at .01 was 136 between a minimum of .24 and a maximum of 8.74. The median percentage of discrepancy for the multivariate mixed model test with nominal alpha set at .05 was 1.23 between a minimum of 0 and a maximum of 5.84. These results also compare favorably with the Monte Carlo results of Olson (1973).

The Mauchly-Thomas chi-square test. The Monte Carlo critical values, and the tabled critical values, for the Mauchly-Thomas Chi-square statistic Table 15 are located in Tables 15 and 16. Table 15 contains the critical values for those Monte Carlo problems where p was set at 2, and Table 16 contains the critical values for those Monte Carlo problems where p was set at 3.

In all cases, the Monte Carlo critical values for the Mauchly-Thomas Chi-square statistic exceeded their tabled critical value counterpart by a wide margin. The median percentage of discrepancy for the Mauchly-Thomas Chi-square test with nominal alpha set at .01 was 36.29 between a minimum of 13.43 and a maximum of 96.83. The median percentage of discrepancy for the Mauchly-Thomas Chi-square test with nominal alpha set at .05 was 35.11 between a minimum of 122 and a maximum of 88.68.

Discussion

The Doubly Multivariate Test

Because the doubly multivariate F test does not assume multivariate sphericity, it was expected to demonstrate appropriate Type I error rates. In this context, appropriate Type I error rate refers to a satisfactory approximation of the nominal and actual alpha levels. The fact that it did maintain appropriate Type I error control across all levels of the various independent variables provides evidence for the precision of the Monte Carlo procedures.

The Multivariate Mixed Model Test

As expected, the multivariate mixed model test demonstrated appropriate Type I error rates when its assumptions were satisfied. Further, under a mild violation of the multivariate sphericity assumption (i.e., ϵ = .95 in each of the dependent variables), the multivariate mixed model test maintained adequate Type I error control. The liberalness of the multivariate mixed model test under a moderate departure from multivariate sphericity is similar in magnitude to the liberalness of the mixed model test under a departure from sphericity when p = 1. As a result, the data analyst using a repeated measures design with multiple measures per occasion runs an increased risk of increased Type I error if the multivariate sphericity assumption is not, or is nearly not, satisfied.

Unfortunately, the data analyst in this situation does not have a modified form of the multivariate mixed model test available to control Type I error under violations of the multivariate sphericity assumption. Neither a multivariate analogue of the correction factor ϵ, which could be applied to the multivariate mixed model degrees of freedom, or a conservative form of the multivariate mixed model test, have been defined. It would seem that the derivation of a correction factor would involve the quadratic form of vectors, as well as ratios of central Wisart distributions.

The Mauchly-Thomas Chi-square Test

The most dramatic result of this study is the failure of the Mauchly-Thomas Chi-square test to demonstrate a satisfactory sampling distribution under the null hypothesis. Thomas (personal communication, November 6, 1984) attributes the elevated Type I error rates for the Mauchly-Thomas Chi-square under its null hypothesis to the fact that the test is asymptotic and approaches its true test size only as n approaches infinity. This is reflected in the Monte Carlo results as a decrease in the actual alpha levels as n increases. Further, it explains the elevated Monte Carlo critical values for the Mauchly-Thomas Chi-square statistic. As sample size increases, the Monte Carlo critical values should approximate the corresponding critical tables values. Additional analyses show results which follow the expected trend, however, the size of the n's required to achieve a satisfactory sampling distribution is prohibitively large, as many as several hundred.

Thomas (personal communication, November 6, 1984) suggests three approaches for improving this situation. The following discussion of these three approaches depends heavily upon Thomas' remarks. In the first approach, Thomas suggests replacing n in in Equation 20 with the value (n-g), where g is the number of groups. In the single group case, this replacement value is (n-1). This modification of the Chi-square is offered by Thomas to improve, not correct, the Type I error rate problem. As Thomas notes, this modification is similar to a modification for a likelihood ratio criteria described by Anderson (1958, p. 249). This modified test was examined in the same fashion as was the original test. The correction did decrease the Type I error rate, but not satisfactorily (Robey, 1985).

Thomas' second suggestion for improving the Type I error problem is to determine a multiplier, C, for the likelihood ratio criteria, ω, in the expression $-2\ln(\omega)$. This expression is an equivalent form of Equation 20. Here, is given by

$$\omega = \frac{\left| E^* \right|^{n/2}}{\left| \sum_{i=1}^{K-1} \frac{E_i}{(k-1)} \right|^{n(k-1)/2}} \qquad (22)$$

Here, E* is the doubly multivariate sums of squares and cross products matrix (SSCP) given by Equation 14, and the E_i are the p x p SSCP matrices for the dependent variables at each of the k occasions. The multiplier, C, would be a correction factor which would cause the first two moments of ω to agree with a Chi-square distribution on an appropriate number of degrees of freedom. Based upon the correction factor described by Morrison (1967, p. 153) for Box's M test, Thomas speculates that reasonable Type I error control might be achieved with this correction factor and an n (per group) of twenty or more.

The third approach suggested by Thomas would be to determine the exact distribution of ω. This would parallel the work of Nagarsenker and Pillai (1973), among others, in determining the exact distribution of w in Equation 10. Ultimately, one would prefer this solution.

Statistical Power of the Multivariate F Tests

Davidson (1972) showed that the manner in which the mean differences among the k occasions relate to the structure of a non-spherical variance-covariance matrix can dramatically change the statistical powers of the multivariate and the mixed model tests. A similar power differential can also be observed in the case where $p > 1$. Consider the three data sets in Table 18. The research design for each of these data sets is a repeated measures design where $k = 3$ and $p = 2$. Each of these data sets share the variance-covariance matrix, and the orthonormally transformed SSCP matrix, found in Table 19. As a result, the Mauchly-Thomas Chi-square for all three data sets are equal.

All three data sets represent a situation where relatively large variances were obtained at the first occasion, and smaller variances were obtained in the occasions which follow in time. In real terms, this could reflect a population with some ameliorating clinical disorder. That is, as some treatment is imposed over time, the subjects become more homogeneous.

In data set A, the measures obtained on the second and third occasions are fairly stable, while the mean scores on both of these occasions are substantially different from the mean scores obtained on the first occasion. In this situation, the subjects respond to treatment and then plateau as they become more homogeneous. The multivariate mixed model F test in this case is 5.79 on 4 and 34 degrees of freedom ($p = .0012$). However, the doubly multivariate F test does not detect the mean differences with an F of 75 on 2 and 6 degrees of freedom ($p = .1143$). If a data analyst did not conduct the multivariate mixed model test on the basis of the Mauchly-Thomas Chi-square statistic (Chi-square = 50.02; df = 7.5, $p \leq .005$), the treatment effect could be overlooked. Currently, we are examining a conservative multivariate mixed model test which can detect these differences with acceptable, albeit conservative, Type I error control.

Data sets B and C represent the multivariate extension of Davidson's (1972, p. 451) 'small but reliable' treatment effect. In both cases, small treatment effects are present between occasions 2 and 3 where the variances are relatively small. The multivariate mixed model F test is not at all sensitive to these treatment effects.

Data set B could represent a situation where some treatment is introduced between periods of no treatment for base line and extinction. Here, the multivariate mixed model F test is 0.62 (df = 4,34; $p = .6496$), while the doubly multivariate F test equals 25.02 (df = 4,6; $p = .0007$).

In data set C, the measures increase on the third occasion after remaining stable on the first two. In this case, the multivariate mixed model F test equals .74 (df = 4,34; p =

.5723), and the multivariate mixed model F equals 14.53 (df = 4,6; p = .003).

Two additional aspects of these data sets represent multivariate extensions of Davidson's (1972) findings. First, combinations of the treatment effects found in the three data sets are possible, and are probably likely in practice. The second extension concerns the relative power of the two tests. As seen in Cases B and C, the multivariate mixed model F test may completely miss a treatment effect which can be detected by the doubly multivariate F test. However, although the doubly multivariate mixed model F test does not achieve significance in Case A, it is not completely blind to the effect. In fact, with more subjects, the effect would be detected.

Data Analysis Strategies

The data analysis strategies in this section were developed from a perspective which relates hypothesis testing and inferential statistics by means of an exploratory - confirmatory continuum. The questions a researcher asks in an exploratory phase of research are more general in nature. These questions are best answered by omnibus, or overall, tests. In a confirmatory phase of research, a researcher is in a position to ask specific questions. Answering these questions may require the application of *a priori* contrasts. From this perspective, then, two data analysis strategies are offered; one for exploratory research, and one for confirmatory research.

Exploratory phase of research. In an exploratory stage, the practitioner faced with analyzing data from a single group repeated measures design with multiple measures per occasions must make two difficult decisions. The first decision concerns the preliminary testing for multivariate sphericity. Based upon the results of this study, applying the Mauchly-Thomas Chi-square test seems appropriate only with a large sample size, i.e., 100 to 200 or more, depending upon the size of k. If a preliminary test of multivariate sphericity is desired with smaller sample sizes, certainly the modified Mauchly-Thomas Chi-square test should be used. However, with smaller sample sizes, unless the multivariate sphericity assumption is known to be true, the practitioner might just as well forego a preliminary test of multivariate sphericity and assume a departure from multivariate sphericity exists in the data.

Secondly, the practitioner must choose a test statistic to detect treatment effects. If the multivariate sphericity assumption is known to be true, the multivariate mixed model test is indicated considering its greater number of denominator degrees of freedom. Since the doubly multivariate test and the multivariate mixed model tests may be sensitive to different patterns of treatment effects under a departure from multivariate sphericity, both the doubly multivariate and the multivariate mixed model tests should be conducted. Obviously, in the absence of a corrected, or a conservative, multivariate mixed model test, the situation remains difficult. The importance of maximizing sample size so that the doubly multivariate test might detect treatment effects of the magnitude which the experimenter considers important, must not be overlooked.

Barcikowski and Robey (1984a, 1984b) have suggested conducting both the multivariate and the corrected mixed model test for the $p = 1$ case under a departure from sphericity. In order to maintain Type I error control, they suggested splitting the a priori alpha level equally between the two tests. This protective measure seems appropriate at the $p > 1$ level as well.

Confirmatory stage of research. In a confirmatory stage, it seems very reasonable to extend Rogan, Keselman and Mendoza's (1979) recommendation for the $p = 1$ case that the practitioner conduct a priori cell comparison contrasts instead of an omnibus test. At the multiple measure per occasion level, a priori contrasts are appealing for two reasons. First, using a partitioned error for a multivariate mixed model contrast causes the multivariate sphericity assumption to be trivially satisfied. Therefore, the need for conducting a preliminary test of the multivariate sphericity assumption is eliminated. Second, the F statistic obtained on a multivariate mixed model contrast, using a partitioned error, is the same as that F obtained on a doubly multivariate contrast. Thus, the choice of a test statistic is no longer an issue. Type I error for this approach should be controlled by setting an a priori familywise alpha level.

This approach is recommended for the researcher who is in a position to formulate a set of specific hypotheses contrasting various levels of the independent variable. It is not recommended as a substitute for the omnibus test through the examination of all possible pair comparisons.

REFERENCES

Anderson, T. W. (1958). An introduction to multivariate statistical analysis. New York: John Wiley & Sons.

Barcikowski, R. S., & Robey, R. R. (April, 1984a). Exploratory repeated measures analysis. A paper presented before the annual convention of the American Educational Research Association, New Orleans.

Barcikowski, R. S., & Robey, R. R. (1984b). Decisions in single group repeated measures analysis: statistical tests and three computer packages. The American Statistician, 38, 148-150.

Bock, R. D. (1975). Multivariate statistical methods in behavioral research. New York: McGraw-Hill, Inc.

Boik, R. J. (1981). A priori tests in repeated measures designs: effects of non-sphericity. Psychometrika, 46, 241-255.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality variance and of correlation between errors in the two-way classification. Annals of Mathematical Statistics, 25, 484-498.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences, (2nd·ed.). New York: Academic Press.

Davidson, M. L. (1972). Univariate versus multivariate tests in repeated measures experiments. Psychological Bulletin, 77, 446-452.

Eisenhart, C. (1974). The assumptions underlying the analysis of variance. Biometrics, 3, 1-21.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. Annals of Mathematical Statistics, 29, 885-891.

Halperin, S. (1976, April). Design of Monte Carlo stidies. Paper presented before the annual meeting of the American Educational Research Association, San Francisco.

Huynh, H. (1978). Some approximate tests for repeated measurement designs. Psychometrika, 43, 161-175.

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurement designs have exact F-distributions. Journal of the American Statistical Association, 65, 1582-1589.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. Journal of Educational Statistics, 1, 69-82.

Huynh, H., & Mandeville, G. K. (1979). Validity conditions in repeated measures designs. Psychological Bulletin, 86, 964-973.

Imhof, J. P. (1962). Testing the hypothesis of no fixed main-effects in Scheffe's mixed model. The Annals of Mathematical Statistics, 33, 1085-1095.

International Mathematical and Statistical Libraries, Inc. (1982). The IMSL library (Vol. 2). Houston, TX: Author.

Jensen, D. R. (1982). Efficiency and robustness in the use of repeated measurements. Biometrics, 38, 813-825.

Keppel, G. (1982). Design and analysis: a researchers handbook, (2nd ed.). Englewood Cliff, NJ: Prentice-Hall, Inc.

Keselman, H. J., & Rogan, J. C. (1980). Repeated measures F tests and psychophysiological research: controlling the number of false positives. Psychophysiology, 17, 499-503.

Keselman, H. J., Rogan, J. C., Mendoza, J. L., & Breen, L. J. (1980). Testing the validity conditions of repeated measures F tests. Psychological Bulletin, 87, 479-481.

Kirk, R. E. (1982). Experimental design: procedures for the behavioral sciences, (2nd ed.). Belmont, CA: Brooks/Cole Publishing Co.

Lindquist, E. F. (1953). Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin Co.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. Annals of Mathematical Statistics, 11, 204-209.

McHugh, R. B., Sivanich, G., & Geisser, S. (1961). On the evaluation of personality changes as measured by psychometric test profiles. Psychological Reports, 9, 335-344.

Mendoza, J. L., Toothaker, L. E., & Nicewander, W. A. (1974). A Monte Carlo comparison of the univariate and multivariate methods for the groups by trials repeated measures design. Multivariate Behavioral Research, 9, 165-178.

Morrison, D. F. (1967). Multivariate statistical methods. New York: McGraw-Hill, Inc.

Nargarsenker, B. N., & Pillai, K. C. S. (1973). The distribution of the sphericity criterion. Journal of Multivariate Analysis, 3, 226-235.

Olson, C. L. (1973). A Monte Carlo Investigation of the Robustness of Multivariate Analysis of Variance. Unpublished doctoral dissertation, University of Toronto.

Rich, C. E. (1983). Repeated measures designs, in R. S. Barcikowski (Ed.), Computer packages and research design, Volume 3: SPSS ans SPSSX (pp. 567-710). Lanham, MD: University Press of America.

Robey, R. R. (1985). A Monte Carlo investigation of Type I error in the single group repeated measures design with multiple measures per occasion. Unpublished doctoral dissertation, Ohio University.

Rogan, J. C., Keselman, H. J., & Mendoza, J. L. (1979). Analysis of repeated measures. British Journal of Mathematical and Statistical Psychology, 32, 269-286.

Rouanet, H., & Lépine, D. (1970). Comparison between treatments in a repeated-meausrement design: ANOVA and multivariate methods. The British Journal of Mathematical and Statistical Psychology, 23, 147-163.

Scheffé, H. (1959). The analysis of variance. New York: John Wiley and Sons.

Thomas, D. R. (1983). Univariate repeated measures techniques applied to multivariate data. Psychometrika, 48, 451-462.

Timm, N. H. (1975). Multivariate analysis with applications in education and psychology. Monterey, CA: Brooks/Cole Publishing Co.

Timm, N. H. (1980). Multivariate analysis of repeated measurements, in P. R. Krishnaiah (Ed.), Handbook of statistics. (Vol. 1, pp. 41-87). Amsterdam: North-Holland.

Timm, N. H., & Carlson, J. E. (December, 1973). Multivariate analysis of nonorthogonal experimental designs using a multivariate full rank model. Paper presented before the annual meeting of the American Statistical Society, New York.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. Biometrika, 24, 471-494.

Winer, B. J. (1971). Statistical principles in experimental design, (2nd ed.). New York: McGraw-Hill, Inc.

Appendix

This appendix contains tables of actual alpha levels, as well as tables of Monte Carlo critical values for all three multivariate tests in each Monte Carlo problem. In addition, this appendix contains tables which relate to the example data sets.

Table 1

Actual S imple Sizes for each Combination of k, p and n

	k=3		k=5		k=7	
n	p=2	p=3	p=2	p=3	p=2	p=3
n_1	7	11	15	9	15	21
n_2	14	18	22	16	22	28
n_3	24	28	32	26	32	38

Note. The three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$).

Table 2

Actual Alpha Values of the Doubly Multivariate F Statistic Under Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.013	.008	.009	.015	.009	.010
		.054	.050	.046	.059	.055	.052
r_1	n_2	.013	.010	.008	.007	.006	.009
		.056	.046	.040	.053	.043	.047
	n_3	.010	.010	.013	.016	.007	.013
		.054	.049	.047	.059	.048	.044
	n_1	.009	.007	.011	.014	.007	.013
		.044	.045	.048	.048	.049	.053
r_2	n_2	.009	.014	.010	.009	.010	.009
		.048	.056	.048	.046	.049	.044
	n_3	.010	.011	.007	.011	.011	.011
		.054	.043	.045	.044	.046	.051
	n_1	.011	.009	.011	.011	.011	.013
		.050	.049	.049	.046	.054	.057
r_3	n_2	.007	.012	.010	.009	.011	.009
		.052	.049	.053	.047	.054	.048
	n_3	.007	.010	.009	.008	.008	.013
		.046	.057	.047	.052	.042	.055

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq .01$) departure from the nominal alpha.

Table 3

Actual Alpha Values of the Doubly Multivariate F Statistic Under a Mild Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.009 .045	.011 .046	.008 .039	.013 .053	.014 .054	.015 .058
r_1	n_2	.013 .044	.013 .053	.009 .047	.009 .053	.007 .052	.007 .051
	n_3	.008 .052	.010 .044	.005* .050	.010 .050	.009 .050	.007 .054
	n_1	.009 .049	.008 .047	.008 .046	.011 .050	.014 .055	.014 .051
r_2	n_2	.010 .052	.009 .048	.013 .054	.006 .044	.009 .042	.013 .050
	n_3	.013 .051	.009 .052	.011 .048	.008 .047	.010 .048	.007 .041
	n_1	.006 .044	.008 .044	.013 .054	.011 .059	.010 .054	.011 .054
r_3	n_2	.010 .046	.009 .048	.009 .041	.011 .050	.009 .052	.010 .051
	n_3	.012 .048	.009 .048	.014 .060	.009 .047	.012 .049	.006 .057

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq .01$) departure from the nominal alpha.

Table 4

Actual Alpha Values of the Doubly Multivariate F Statistic Under a Moderate Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.009	.008	.008	.010	.005	.013
		.040	.042	.045	.050	.042	.044
r_1	n_2	.008	.010	.010	.010	.012	.011
		.056	.052	.049	.046	.058	.048
	n_3	.013	.009	.010	.011	.015	.009
		.049	.038*	.044	.051	.059	.055
	n_1	.014	.011	.012	.010	.009	.009
		.048	.048	.049	.047	.044	.044
r_2	n_2	.009	.010	.013	.009	.008	.011
		.055	.049	.049	.048	.043	.046
	n_3	.012	.011	.013	.010	.010	.012
		.051	.048	.051	.050	.046	.044
	n_1	.014	.011	.009	.014	.013	.009
		.052	.049	.045	.053	.050	.053
r_3	n_2	.012	.010	.007	.008	.011	.008
		.052	.051	.052	.045	.054	.054
	n_3	.011	.011	.011	.018*	.008	.013
		.058	.051	.050	.058	.049	.059

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes (n_1 = (k-1)p + 3, n_2 = (k-1)p + 10, n_3 = (k-1)p + 20), and the three correlations among the dependent measures across occasions (r_1 = .2, r_2 = .5, r_3 = .8). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq$ .01) departure from the nominal alpha.

Table 5

Actual Alpha Values of the Mixed Model Multivariate F Statistic Under Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.008	.012	.011	.009	.007	.013
		.056	.056	.045	.040	.041	.053
r_1	n_2	.012	.008	.006	.008	.008	.011
		.054	.053	.044	.055	.053	.053
	n_3	.012	.007	.013	.011	.010	.012
		.050	.049	.053	.056	.047	.059
	n_1	.008	.004*	.013	.008	.010	.011
		.049	.044	.054	.047	.053	.043
r_2	n_2	.009	.013	.008	.007	.007	.012
		.050	.046	.057	.044	.048	.053
	n_3	.011	.009	.012	.006	.009	.009
		.046	.053	.049	.044	.051	.049
	n_1	.006	.009	.012	.012	.008	.008
		.053	.054	.055	.049	.051	.047
r_3	n_2	.007	.007	.013	.011	.011	.008
		.049	.048	.059	.045	.050	.049
	n_3	.007	.015	.011	.014	.015	.012
		.046	.059	.058	.047	.044	.044

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq .01$) departure from the nominal alpha.

Table 6

Actual Alpha Values of the Mixed Model Multivariate F Statistic Under a Mild Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.012	.013	.014	.008	.010	.012
		.049	.051	.051	.051	.046	.054
r_1	n_2	.012	.011	.012	.011	.008	.013
		.048	.055	.051	.052	.049	.054
	n_3	.010	.008	.014	.013	.010	.010
		.053	.054	.051	.056	.050	.060
	n_1	.013	.009	.011	.013	.009	.014
		.050	.052	.056	.054	.049	.054
r_2	n_2	.012	.016	.012	.012	.011	.009
		.048	.071*	.056	.055	.054	.048
	n_3	.013	.012	.012	.009	.011	.010
		.057	.052	.049	.050	.063	.052
	n_1	.015	.009	.012	.014	.014	.012
		.043	.051	.052	.056	.063	.057
r_3	n_2	.013	.014	.012	.012	.012	.012
		.063	.045	.053	.056	.062	.056
	n_3	.015	.014	.012	.013	.014	.010
		.055	.051	.059	.051	.056	.054

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq .01$) departure from the nominal alpha.

Table 7

Actual Alpha Values of the Mixed Model Multivariate F Statistic Under a Moderate Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.022*	.019*	.022*	.013	.023*	.021*
		.067*	.063	.075*	.069*	.071*	.067*
r_1	n_2	.020*	.021*	.022*	.021*	.022*	.015
		.074*	.062	.073*	.065*	.079*	.061
	n_3	.020*	.021*	.021*	.021*	.026*	.021*
		.070*	.073*	.069*	.075*	.084*	.071*
	n_1	.017*	.019*	.022*	.018*	.026*	.028*
		.058	.071*	.069*	.076*	.072*	.075*
r_2	n_2	.016	.026*	.021*	.011	.024*	.019*
		.061	.079*	.068*	.062	.072*	.070*
	n_3	.022*	.024*	.023*	.016	.021*	.022*
		.064*	.066*	.066*	.069*	.068*	.072*
	n_1	.024*	.018*	.014	.022*	.018*	.026*
		.074*	.075*	.065*	.081*	.070*	.080*
r_3	n_2	.020*	.022*	.016	.025*	.019*	.028*
		.067*	.071*	.069*	.093*	.070*	.077*
	n_3	.026*	.016	.023*	.026*	.018*	.033*
		.070*	.073*	.071*	.076*	.068*	.086*

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes (n_1 = (k-1)p + 3, n_2 = (k-1)p + 10, n_3 = (k-1)p + 20), and the three correlations among the dependent measures across occasions (r_1 = .2, r_2 = .5, r_3 = .8). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq$ .01) departure from the nominal alpha.

Table 8

Actual Alpha Values of the Mauchly-Thomas Chi-square Statistic Under Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.174*	.349*	.532*	.866*	.873*	.994*
		.352*	.547*	.737*	.949*	.954*	.998*
r_1	n_2	.045*	.081*	.144*	.387*	.391*	.837*
		.133*	.211*	.320*	.612*	.634*	.946*
	n_3	.026*	.029*	.063*	.145*	.157*	.464*
		.087*	.128*	.173*	.328*	.353*	.699*
	n_1	.173*	.332*	.526*	.880*	.873*	.997*
		.341*	.552*	.738*	.959*	.950*	1.000*
r_2	n_2	.043*	.089*	.139*	.370*	.376*	.833*
		.145*	.215*	.312*	.609*	.612*	.943*
	n_3	.030*	.043*	.055*	.149*	.152*	.470*
		.094*	.138*	.174*	.358*	.333*	.714*
	n_1	.167*	.339*	.528*	.856*	.872*	.996*
		.342*	.555*	.721*	.948*	.959*	.999*
r_3	n_2	.049*	.076*	.151*	.373*	.373*	.824*
		.154*	.212*	.336*	.605*	.608*	.933*
	n_3	.027*	.034*	.060*	.165*	.143*	.483*
		.093*	.130*	.173*	.365*	.335*	.712*

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom). An asterisk indicates a significant ($p \geq .01$) departure from the nominal alpha.

Table 9

Proportions of Rejection for the Mauchly-Thomas Chi-square Statistic Under a Mild Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.185	.377	.595	.918	.932	1.000
		.366	.585	.790	.974	.981	1.000
r_1	n_2	.077	.137	.241	.573	.613	.943
		.212	.307	.466	.778	.815	.988
	n_3	.078	.118	.183	.384	.452	.842
		.224	.290	.382	.647	.703	.947
	n_1	.192	.378	.608	.919	.912	1.000
		.381	.600	.794	.978	.974	1.000
r_2	n_2	.073	.134	.260	.576	.615	.955
		.195	.314	.478	.786	.808	.988
	n_3	.078	.111	.186	.382	.419	.837
		.213	.274	.392	.648	.675	.948
	n_1	.193	.392	.600	.929	.929	.998
		.384	.603	.792	.982	.977	1.000
r_3	n_2	.073	.144	.247	.578	.601	.950
		.208	.340	.462	.780	.812	.985
	n_3	.078	.123	.197	.412	.451	.819
		.229	.288	.417	.652	.688	.946

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes ($n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, $n_3 = (k-1)p + 20$), and the three correlations among the dependent measures across occasions ($r_1 = .2$, $r_2 = .5$, $r_3 = .8$). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom).

Table 10

Proportions of Rejection for the Mauchly-Thomas Chi-square Statistic Under a Moderate Departure from Multivariate Sphericity

		k=3		k=5		k=7	
rho	n	p=2	p=3	p=2	p=3	p=2	p=3
	n_1	.369	.654	.848	.993	.994	1.000
		.599	.839	.937	.998	.998	1.000
r_1	n_2	.483	.726	.810	.986	.992	1.000
		.715	.878	.930	.998	.998	1.000
	n_3	.809	.924	.936	.996	.999	1.000
		.929	.982	.982	1.000	1.000	1.000
	n_1	.379	.659	.860	.993	.995	1.000
		.612	.843	.952	.998	.999	1.000
r_2	n_2	.505	.741	.807	.981	.995	1.000
		.726	.899	.936	.996	.998	1.000
	n_3	.810	.927	.943	.996	.999	1.000
		.939	.985	.984	.999	1.000	1.000
	n_1	.385	.660	.856	.993	.995	1.000
		.591	.840	.951	.999	1.000	1.000
r_3	n_2	.499	.712	.812	.984	.996	1.000
		.715	.887	.937	1.000	1.000	1.000
	n_3	.800	.932	.947	.994	.999	1.000
		.931	.985	.990	.999	1.000	1.000

Note. Monte Carlo problems are defined by combinations of the k occasions, the p measures, the three sample sizes (n_1 = (k-1)p + 3, n_2 = (k-1)p + 10, n_3 = (k-1)p + 20), and the three correlations among the dependent measures across occasions (r_1 = .2, r_2 = .5, r_3 = .8). Note that the double entries for each Monte Carlo problem represent nominal alpha at .01 (top), and at .05 (bottom).

Table 11

Tabled Critical Values and Monte Carlo Critical Values for the Doubly Multivariate F Test Under Multivariate Sphericity with Two Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	28.71	33.04	28.05	29.56
		9.12	9.79	8.31	9.25
3	n_2	5.99	6.57	5.70	5.72
		3.48	3.62	3.38	3.52
3	n_3	4.43	4.43	4.43	4.26
		2.87	2.95	2.92	2.80
5	n_1	27.50	24.53	28.66	30.64
		8.55	8.52	8.51	8.61
5	n_2	5.06	4.76	5.05	4.76
		3.07	2.82	3.04	3.18
5	n_3	3.56	3.70	3.33	3.55
		2.45	2.37	2.38	2.39
7	n_1	27.03	25.35	22.16	30.76
		8.74	9.32	8.58	9.26
7	n_2	4.71	4.29	4.64	4.88
		2.91	2.77	2.90	2.97
7	n_3	3.23	2.98	3.27	3.12
		2.28	2.25	2.22	2.17

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 12

Tabled Critical Values and Monte Carlo Critical Values for the Doubly Multivariate F Test Under Multivariate Sphericity with Three Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	27.91	25.16	23.67	27.63
		8.94	8.94	8.27	8.72
3	n_2	5.39	5.30	5.59	5.54
		3.22	3.15	3.37	3.18
3	n_3	3.87	3.84	3.92	3.82
		2.60	2.57	2.49	2.65
5	n_1	27.03	35.12	33.48	27.34
		8.74	10.14	8.49	8.19
5	n_2	4.71	4.35	4.63	4.56
		2.91	3.02	2.84	2.88
5	n_3	3.23	3.70	3.24	3.08
		2.28	2.37	2.23	2.30
7	n_1	26.77	26.43	37.41	29.98
		8.67	9.23	8.86	9.28
7	n_2	4.46	4.40	4.32	4.37
		2.80	2.74	2.72	2.76
7	n_3	2.99	3.04	3.03	3.24
		2.15	2.11	2.17	2.19

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 13

Tabled Critical Values and Monte Carlo Critical Values for the Multivariate Mixed Model F Test Under Multivariate Sphericity with Two Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	4.14	4.24	4.10	4.15
		2.74	2.90	2.79	2.90
3	n_2	3.69	3.86	3.65	3.42
		2.54	2.62	2.56	2.52
3	n_3	3.53	3.73	3.68	3.39
		2.47	2.48	2.45	2.40
5	n_1	2.73	2.78	2.84	2.81
		2.05	2.01	2.12	2.13
5	n_2	2.64	2.49	2.59	2.77
		2.00	1.96	2.06	2.07
5	n_3	2.59	2.69	2.65	2.60
		1.98	2.01	1.97	2.06
7	n_1	2.29	2.17	2.28	2.23
		1.81	1.74	1.82	1.82
7	n_2	2.25	2.19	2.19	2.28
		1.79	1.80	1.78	1.78
7	n_3	2.23	2.22	2.19	2.35
		1.78	1.77	1.78	1.74

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 14

Tabled Critical Values and Monte Carlo Critical Values for the Multivariate Mixed Model F Test Under Multivariate Sphericity with Three Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	3.43	3.58	3.13	3.52
		2.40	2.54	2.38	2.49
3	n_2	3.12	3.04	3.27	2.97
		2.25	2.30	2.22	2.24
3	n_3	2.99	2.96	2.92	3.10
		2.19	2.18	2.21	2.24
5	n_1	2.30	2.23	2.28	2.34
		1.82	1.77	1.81	1.81
5	n_2	2.26	2.22	2.21	2.28
		1.79	1.83	1.75	1.77
5	n_3	2.24	2.25	2.16	2.33
		1.78	1.80	1.74	1.75
7	n_1	1.99	2.05	2.05	1.92
		1.63	1.66	1.61	1.62
7	n_2	1.97	1.98	2.03	1.93
		1.63	1.64	1.65	1.62
7	n_3	1.96	2.09	1.95	2.03
		1.62	1.64	1.61	1.60

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 15

Tabled Critical Values and Monte Carlo Critical Values for the Mauchly-Thomas Chi-square Test Under Multivariate Sphericity with Two Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	18.47	34.32	34.29	34.30
		14.07	25.28	25.32	26.00
3	n_2	18.47	24.93	23.80	23.67
		14.07	17.91	17.62	18.35
3	n_3	18.47	20.95	20.96	22.08
		14.07	15.80	16.39	16.33
5	n_1	54.80	105.67	96.78	98.80
		47.41	84.04	84.72	84.01
5	n_2	54.80	72.65	71.56	72.53
		47.41	62.61	62.90	63.43
5	n_3	54.80	63.60	66.68	63.14
		47.41	56.50	55.14	55.66
7	n_1	106.41	176.39	167.30	165.39
		96.21	166.24	170.22	160.87
7	n_2	106.41	147.63	149.06	145.00
		96.21	130.30	132.55	130.82
7	n_3	106.41	130.16	128.47	127.13
		96.21	118.20	116.93	115.68

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 16

Tabled Critical Values and Monte Carlo Critical Values for the Mauchly-Thomas Chi-square Test Under Multivariate Sphericity with Three Dependent Variables per Occasion

		Tabled	Monte Carlo Critical Values		
k	n	Value	r = .2	r = .5	r = .8
3	n_1	30.61	60.25	57.57	59.34
		25.00	47.17	46.23	47.02
3	n_2	30.61	39.48	42.29	39.53
		25.00	33.10	33.70	33.03
3	n_3	30.61	34.96	36.30	36.93
		25.00	28.45	29.84	29.21
5	n_1	102.83	169.15	159.84	159.85
		92.80	157.00	158.20	157.96
5	n_2	102.83	143.14	144.86	140.17
		92.80	127.48	128.93	128.56
5	n_3	102.83	123.55	124.65	125.74
		92.80	112.94	112.77	112.93
7	n_1	210.19	326.41	330.61	298.72
		195.97	367.16	315.91	343.11
7	n_2	210.19	289.02	290.93	288.46
		195.97	281.96	295.65	274.80
7	n_3	210.19	266.85	268.15	271.65
		195.97	246.39	250.44	251.63

Note. Sample sizes are defined as $n_1 = (k-1)p + 3$, $n_2 = (k-1)p + 10$, and $n_3 = (k-1)p + 20$. The values of r are the constant correlations among the dependent variables across occasions. The double entries for each Monte Carlo problem represent alpha levels of .01 (top) and .05 (bottom).

Table 17

Monte Carlo Problems Where the Null Hypothesis of the Proportions Test was not Rejected When Applied to the Multivariate Mixed Model Actual Alpha Rates

	Departure from Multivariate Sphericity	Correlation Among p Across k	k	p	n	Nominal Alpha
1.	moderate	.2	3	3	(k-1)p+3	.05
2.	moderate	.2	5	3	(k-1)p+3	.01
3.	moderate	.2	3	3	(k-1)p+10	.05
4.	moderate	.2	7	3	(k-1)p+10	.01
5.	moderate	.2	7	3	(k-1)p+10	.05
6.	moderate	.5	3	2	(k-1)p+3	.05
7.	moderate	.5	3	2	(k-1)p+10	.01
8.	moderate	.5	3	2	(k-1)p+10	.05
9.	moderate	.5	5	3	(k-1)p+10	.01
10.	moderate	.5	5	3	(k-1)p+10	.05
11.	moderate	.5	5	3	(k-1)p+20	.01
12.	moderate	.8	5	2	(k-1)p+10	.01
13.	moderate	.8	5	2	(k-1)p+20	.01
14.	moderate	.8	3	3	(k-1)p+20	.01

Table 18

Three Example Data Sets

	p			p		
Case	k	k	k	k	k	k
	34	58	58	168	169	172
	54	62	62	176	178	179
	26	55	56	149	169	168
	30	58	61	141	173	174
A	16	54	53	132	171	170
	63	60	62	180	179	172
	51	61	59	166	177	178
	47	56	56	137	171	172
	68	64	63	163	175	176
	23	56	53	143	172	168
Mean	41.2	58.4	58.3	155.5	173.4	172.9
	34	37	43	168	148	157
	54	41	47	176	157	164
	26	34	41	149	148	153
	30	37	46	141	152	159
B	16	33	38	132	150	155
	63	39	47	180	158	157
	51	40	44	166	156	163
	47	35	41	137	150	157
	68	43	48	163	154	161
	23	35	38	143	151	153
Mean	41.2	37.4	43.3	155.5	152.4	157.9
	34	42	46	168	152	161
	54	46	50	176	161	168
	26	39	44	149	152	157
	30	42	49	141	156	163
C	16	38	41	132	154	159
	63	44	50	180	162	161
	51	45	47	166	160	167
	47	40	44	137	154	161
	68	48	51	163	158	165
	23	40	41	143	155	157
Mean	41.2	42.4	46.3	155.5	156.4	161.9

Table 19

The Variance-Covariance Matrix and the Orthonormally Transformed SSCP Matrix for the Example Data

317.96	50.69	53.82	220.44	48.24	46.36
	10.71	10.64	42.56	8.93	10.49
		13.79	46.06	9.31	10.59
			294.06	42.11	36.72
				13.38	9.71
(sym)					14.77

1008.45	-582.17	623.80	-343.70
	355.35	-377.82	206.97
		1059.20	-557.26
(sym)			342.73

Note. The variance-covariance matrix is the upper matrix while the orthonormally transformed SSCP matrix is the lower matrix. The variables in the variance covariance matrix are grouped first by p, and ordered by k within p.