

Nonparametric Competitors to the Two-Way ANOVA

Larry E. Toothaker and De Newman
University of Oklahoma

Key words: *two-way designs, ANOVA, nonparametric tests*

The ANOVA F and several nonparametric competitors for two-way designs were compared for empirical α and power. Simulation of 2×2 , 2×4 , and 4×4 designs was done with cell sizes of 5 and 10 when sampling from normal, exponential, and mixed normal distributions. Conservatism of both α and power in the presence of other nonnull effects was seen in the tests due to Puri and Sen (1985) and, to a lesser degree, in the rank transform tests (Conover & Iman, 1981). Tests by McSweeney (1967) and Hettmansperger (1984) had liberal α for some designs and distributions, especially for small n . The ANOVA F suffers from conservative α and power for the mixed normal distribution, but it is generally recommended.

The applied researcher has often expressed concerns about meeting the assumptions of the parametric fixed effects two-way ANOVA. Given the linear model of $Y_{ijk} = \mu + \alpha_j + \beta_k + \alpha\beta_{jk} + e_{ijk}$, the assumptions on the error are that $e_{ijk} \stackrel{d}{\sim} \text{NID}(0, \sigma_e^2)$ for each cell. These concerns about assumptions have been addressed in research on ANOVA, including development of alternatives to the ANOVA and examination of the ANOVA under assumption violations.

Nonparametric competitors to the two-way ANOVA relax the assumptions on e_{ijk} by assuming that they are independent, identically distributed random variables from some continuous distribution. Nonparametrics for the two-way layout include rank tests by Puri and Sen (1985), rank transform tests by Conover and Iman (1981), and aligned rank tests. The latter area has grown out of a suggestion by Hodges and Lehmann (1962) and includes a test for interaction by Mehra and Sen (1969), tests for main effects in an additive model by Mehra and Sarangi (1967) and Sen (1968), tests for interaction in mixed models by Koch (1969) and Koch and Sen (1968), a test for interaction in the presence of main effects by McSweeney (1967), and linear-model-based tests of main effects and interaction by Hettmansperger (1984). Generally, nonparametric methods have gained attention due to their ability to analyze data that is questionable with respect to the normality assumption. These nonparametric competitors to the two-way ANOVA would be desirable alternatives for nonnormal data likely to be found in applied

research (see Micceri, 1989) if they provide adequate control of α and power competitive to that of the F tests in the ANOVA.

The purpose of this article is to present the results of Monte Carlo simulations of four general groups of statistical methods for two-way designs: (a) parametric two-way ANOVA, (b) Puri and Sen (1985) tests, (c) rank transform tests, and (d) aligned rank tests. The latter group includes tests for all three effects from McSweeney (1967), Hettmansperger (1984), and ANOVA F tests on ranks aligned on all other effects possible.

Rank Tests

Puri and Sen Tests

The rank tests given by Puri and Sen (PS) (1985) for linear models can be adopted for tests on main effects and interaction in a two-way layout and can be presented as a function of the proportion of variability due to the desired effect (see Harwell, 1991; Harwell & Serlin, 1990). Given n observations per cell, J rows, and K columns, the steps to do these tests are as follows:

1. Rank all the data from one to total sample size, $N = nJK$.
2. Compute the sums of squares for rows, SS_R , columns, SS_C , interaction, SS_{RC} , and total, SS_T .
3. Compute the statistics $(N - 1)SS_R/SS_T$, $(N - 1)SS_C/SS_T$, and $(N - 1)SS_{RC}/SS_T$.
4. Refer the statistics to chi-square critical values with $df_R = J - 1$, $df_C = K - 1$, and $df_{RC} = (J - 1)(K - 1)$, respectively.

Steps 1 and 2 could be accomplished easily with any major statistical package such as SAS (1990), using PROC RANK to do the ranking and PROC ANOVA to compute the sums of squares. It should be noted that these procedures are considered large-sample tests, because the statistics are only asymptotically chi-square.

Other researchers developed equivalents of the PS rank tests, but Puri and Sen (1985) were the first to present them in a general linear model approach which has broader use than the two-way ANOVA. Reinach (1965) suggested that rank tests using an ANOVA approach would be confounded by the presence of other nonnull effects. Scheirer, Ray, and Hare (1976) gave statistics that are equivalent to the PS statistics for a two-way design and presented Monte Carlo results for small n with a normal distribution. Scheirer et al. (1976) found that the statistics gave adequate control of Type I error rate for main effects and interaction tests, but they investigated only the completely null case. That is, the probability of Type I error was examined for each of the row, column, and interaction tests in the presence of null effects. For a normal distribution, Toothaker and Chang (1980) found that the Scheirer et al. statistics (PS tests) exhibited a dampening effect in the presence of any other nonnull effect. That is, the PS tests became more conservative as the

magnitude of the other effects increased. The conservatism of the PS tests was found for both α and power and for both main effects and interaction tests. This conservatism was found even in the presence of one nonnull effect which would give power of only about .60. Also, Lemmer (1980) presented a statistic which is equivalent to PS tests, and found conservatism of α which became worse as magnitude or number of other effects increased.

Harwell (1991) presented results for the PS tests for small n and found that the PS tests controlled α in the completely null case for five distributions: normal, double-exponential, approximate Cauchy, chi-square with $df = 8$, and chi-square with $df = 2$ (exponential). However, when one nonnull effect, which would give power of about .30 for $n = 5$, is present, PS tests showed conservatism consistent with that found by Toothaker and Chang (1980). This conservatism of the PS tests was most severe for the chi-square distribution with $df = 2$. Power results showed "almost uniformly lower power values than its [PS test] competitors for all distributions" (p. 396), but Harwell (1991) concluded that there was no power decrement or dampening effect on power. Harwell did not observe the dampening effect on power in the PS statistics because he used only one degree of nonnullity for each sample size. However, evidence of the decrement in power is available from his data: when sampling from a normal distribution with one nonnull effect, for $n = 5$, power is .313 for interaction, but, when all effects are nonnull, power is .256. Similarly, for the chi-square distribution with $df = 2$, the power decrement is from .552 to .319. However, the dampening effect is not always present for all PS tests for all distributions. Harwell (1991) also noted a potential problem for PS tests in that the ranking process "can create a serious mismatch between effects in the model purported to underlie the original scores and effects in the ranks of the scores" (p. 384). More will be given on this problem when rank transform tests are discussed.

Rank Transform Tests (RT)

Tests known as rank transforms have been popularized by Conover and Iman (1976, 1981), Iman (1974), and Iman and Conover (1976). The steps to do RT tests in the two-way layout are as follows:

1. Rank all the data from one to total sample size, $N = nJK$.
2. Compute the sums of squares for rows, SS_R , columns, SS_C , interaction, SS_{RC} , and total, SS_T .
3. Compute the ANOVA F statistics $F_R = MS_R/MS_w$, $F_C = MS_C/MS_w$, and $F_{RC} = MS_{RC}/MS_w$.
4. Refer the statistics to F critical values with $df_R = J - 1$, $df_C = K - 1$, and $df_{RC} = (J - 1)(K - 1)$, in the numerators, respectively, and $df_w = JK(n - 1)$ in the denominators.

Again, these steps could easily be accomplished with any major statistical package such as SAS, using PROC RANK to do the ranking and PROC

ANOVA to compute the F ratios. In fact, SAS (1990, p. 493) says that data can be submitted to PROC RANK and then fitted with an ANOVA model using the ANOVA or GLM procedures. It should be noted that RT procedures are considered large-sample tests, because the statistics are only asymptotically distributed as F .

In spite of much positive research on RT tests for two-way layouts (see Hora & Conover, 1984; Hora & Iman, 1988; Iman, 1974; Iman & Conover, 1976; Iman, Hora, & Conover, 1984; Thompson & Ammann, 1989), recent results show RT tests lack control of α in the presence of other nonnull effects (see Sawilowsky, 1990). Fawcett and Salter (1984) found increasing conservatism as the degree of nonnullity of block effects increased in a RT test for treatments in an additive model randomized block design with $n = 1$. In an additive model randomized block design with $n > 1$, Groggel (1987) found that a RT test for treatment evidenced a decrease in power as block effects increased. Akritas (1990) showed that RT tests are confounded due to other nonnull effects. Thus RT tests should not be used to test for main effects in the presence of interaction, nor should RT tests be used to test interaction in the presence of main effects. The following all show that RT tests should not be used for two-way and higher ANOVA layouts: inflation of α for the F for interaction in the presence of nonnull row and column effects in a two-way layout (Blair, Sawilowsky, & Higgins, 1987), inflation of α for any effect and power for nonnull effects in a $2 \times 2 \times 2$ layout (Sawilowsky, Blair, & Higgins, 1989), α and power properties being dependent on how the treatment effects were modeled in a $2 \times 2 \times 2$ layout (Sawilowsky, 1985), and inflation of α for a main effect in the presence of nonnull interaction (Lemmer, 1980). One exception to this warning is the 2×2 layout, where the correct effects seem to be modeled by RT tests (Blair, Sawilowsky, & Higgins, 1987) but where power is lower than the ANOVA on raw data for normal distributions (Sawilowsky, 1989). These results validate concerns expressed by Fligner (1981), "Rank transformation methods can provide quick solutions to many of these problems, but the resulting solutions may not be the best, or may not even be appropriate nonparametric methods." RT tests are included here to further examine these methods for two-way layouts.

Aligned Rank Tests

Given the problems encountered by PS tests and RT tests when other nonnull effects are present, one obvious solution would be to treat other effects as nuisance parameters and remove them from the scores before the ranking and analysis. Indeed, this is the principle on which aligned rank tests are based. For a two-way layout, the general steps to do aligned rank tests are as follows:

1. Estimate the effects on which the aligning is to be done.

2. Subtract the estimated effects from the original scores (align the scores).
3. Rank the aligned scores in the total sample.
4. Compute the test statistic on the ranks of the aligned scores.

The actual effects of the model used to align scores and the eventual test statistic are the basis for the differences that exist in the several forms of aligned rank tests. Variations include different ways to estimate the desired effects—such as, means, trimmed means, medians—and different ways to form the test statistic—such as, using some function of sums of squares, which is asymptotically distributed as χ^2 , or an F statistic, which asymptotically follows an F distribution.

McSweeney (1967) developed a test (M test) for interaction using aligned ranks in the two-way layout with a nonadditive model. After the original scores have been aligned by subtracting estimates of both main effects and after the aligned scores have been ranked, the test statistic for interaction can be computed as $(N - 1)SS_{RC}/SS_T$ and referred to a chi-square critical value with $df_{RC} = (J - 1)(K - 1)$. By analogy, the test for row (column) main effects can be done as follows: Align scores on the estimates of the column (row) and interaction effects; rank the aligned scores, and compute $(N - 1)SS_R/SS_T$ (or $(N - 1)SS_C/SS_T$), which is referred to a chi-square critical value with $df_R = J - 1$ (or $df_C = K - 1$).

Hettmansperger (1984) developed a linear model or regression-based system from which tests (H tests) for a two-way layout could be identified. Harwell (1991) presents these as procedures using ranks on aligned scores, where the aligning is done on all effects in the model that are not being tested. For example, for the interaction test, after the original scores have been aligned by subtracting estimates of both main effects, obtain ranks of the aligned scores, r_{ijk} . Standardize the ranks to have mean zero and variance one by using

$$r'_{ijk} = \left(\frac{r_{ijk}}{N + 1} - .5 \right) (\sqrt{12}), \quad (1)$$

and use r'_{ijk} to compute SS_{RC} . Two different critical values were proposed by Hettmansperger (1984): a chi-square critical value with $df_{RC} = (J - 1)(K - 1)$ and a $(df_{RC})F$ critical value with degrees of freedom df_{RC} and df_w . Preference is given for the latter due to liberal simulation results (cited in Hettmansperger, 1984) when using the chi-square critical value. Similarly, the test for row (column) main effects can be done as follows: Align scores on the estimates of the column (row) and interaction effects; rank the aligned scores; standardize the ranks using Equation 1; use r'_{ijk} to compute SS_R (or SS_C), which is referred to a chi-square critical value with $df_R = J - 1$ (or $df_C = K - 1$). An alternate F -based critical value would be

$(df_R)F$ with degrees of freedom df_R and df_w (or $(df_C)F$ with degrees of freedom df_C and df_w). Note that three separate rankings must be done for McSweeney or Hettmansperger tests.

A final aligned rank test to be considered would be the ANOVA F tests on ranks aligned on all other effects possible (RO tests). The aligning and ranking are the same as those shared by the McSweeney and Hettmansperger tests, but the test statistics differ. After aligning the raw data by subtracting estimates of both main effects and obtaining ranks of all of the aligned data, the statistic for interaction is the ANOVA F on ranks, which is then referred to a F critical value with df_{RC} and df_w . Statistics for the main effects come from two additional rankings, where the data are first aligned by subtracting estimates of the other main effect and the interaction. Then the ANOVA F_R (F_C) on ranks is referred to a F critical value with df_R and df_w (df_C and df_w). These tests could be thought of as rank transform tests after aligning on all other possible effects.

In summary, the PS, RT, M, H, and RO tests for two-factor designs have in common the use of ranks, but they differ with respect to how the ranks are operationalized into a test statistic. The PS and RT tests do not align the observations before ranking, and the PS tests are functions of sums of squares, which are referred to chi-square critical values, while the RT tests compute F s, which are referred to F critical values. M, H, and RO all align on all other effects but compute functions of sums of squares referred to chi-square critical values, sums of squares of standardized ranks referred to $(df)F$ critical values, and F s which are referred to F critical values, respectively.

Method

A computer program was written to perform simulations of the various tests, including the ANOVA F , for 2×2 , 2×4 , and 4×4 designs, for cell sizes of $n = 5$ and $n = 10$, and for normal, exponential, and mixed-normal distributions. Microsoft FORTRAN was used, and the program was run on a Tri-Star 386/33 with a math coprocessor.

Because of similar results for various levels of α in previous studies (see Harwell, 1991; Toothaker & Chang, 1980), only nominal α of .05 was used. One null and two nonnull cases were examined for both main and interaction effects. Nonnull effects were chosen so as to give normal-theory power of approximately .85, called *medium*, and .99, called *large*. Table 1 shows all cases obtained by combining main and interaction effects.

For the normal population, data were generated using the Box and Muller (1958) method, transforming independent unit uniform pseudorandom numbers from a procedure by Chen (1971). For the exponential population, data were generated by a method reported by Lehmann and Bailey (1968): $f(t) = pe^{-pt}$, with $p = 1$, $E(t) = 1/p = 1$, and $\text{var}(t) = 1/p^2 = 1$. Pseudorandom exponential variables were generated by multiplying -1 by the natural loga-

TABLE 1
Cases of main and interaction effects

Case	A main effect	B main effect	Interaction effect
1	Null ^a	Null	Null
2	Null	Null	Medium ^b
3	Null	Null	Large ^c
4	Null	Medium	Null
5	Null	Medium	Medium
6	Null	Medium	Large
7	Null	Large	Null
8	Null	Large	Medium
9	Null	Large	Large
10	Medium	Null	Null
11	Medium	Null	Medium
12	Medium	Null	Large
13	Medium	Medium	Null
14	Medium	Medium	Medium
15	Medium	Medium	Large
16	Medium	Large	Null
17	Medium	Large	Medium
18	Medium	Large	Large
19	Large	Null	Null
20	Large	Null	Medium
21	Large	Null	Large
22	Large	Medium	Null
23	Large	Medium	Medium
24	Large	Medium	Large
25	Large	Large	Null
26	Large	Large	Medium
27	Large	Large	Large

^a Null = all effects are equal to zero, ^b Medium = effects are selected so that power will be about .85 for the normal distribution, ^c Large = effects are selected so that power will be about .99 for the normal distribution.

rithm of unit uniform pseudorandom numbers (Chen, 1971). The exponential variates were then scaled to have means determined by the null or nonnull conditions and variance of one. For the mixed normal population, data were generated by generating 95% of the scores from $N(0, 1)$ and 5% of the scores from $N(22, 100)$ using the Box and Muller (1958) and Chen (1971) methods. These variates, with mean of 1.1 and variance of 8.375, were then scaled to have the mean determined by the null or nonnull conditions and variance of one. For any tests requiring aligning, estimates of aligned effects were computed using appropriate sample means.

Empirical α and power were obtained over 5,000 replications, but the main effect information was collapsed for those cases where the degrees of freedom and the process of ranking/aligning were the same for both main effects. Thus, the values in the tables for those cases are averages of two proportions obtained over 5,000 replications. Note that all 27 cases in Table 1 were examined for the same 5,000 replications, thus controlling sampling variability between cases for each design, sample size, and distribution combination.

Results

Because the sampling variability of the empirical estimates was small, $\sigma_p = .0030822$ for $\alpha = .05$ and 5,000 replications, arbitrary standards were used to decide when a statistic was evidencing lack of α robustness. Values of empirical α that deviated $\pm .01$ from $.05$ were called slightly liberal or slightly conservative, and values that deviated $\pm .02$ were called liberal or conservative. Thus, a value of empirical α of $.057$ was more than 2 standard deviations from $.05$ but was deemed not indicative of lack of robustness. Power robustness was determined by comparison of values for the various statistical procedures—that is, a method was called conservative if it showed power which was substantially less than that for the other methods.

Normal Distribution

Empirical α results for $n = 5$ showed (see Table 2) M tests to be slightly liberal or liberal for all designs. H tests were slightly liberal or liberal for the main effect tests in the larger two designs and for the interaction test in the smaller two designs. F tests were not liberal or conservative, and RO tests were slightly liberal only for interaction in the 2×2 design. PS tests showed conservatism in the presence of other nonnull effects which worsened as the number of nonnull effects or degree of nonnullity increased. For example, in the 2×2 design, in the presence of the other main effect with medium power (Cases 4 and 10), empirical α for the PS main effect test was $.0177$, but, in the presence of the other two effects with large power (Cases 9 and 21), the main effect had an empirical α of $.0002$. Similar results were seen for the PS interaction test. RT tests showed conservatism in the 2×2 design only if both other effects were nonnull. Power results showed PS tests with conservatism in the presence of other nonnull effects that dramatically worsened as the number of nonnull effects or degree of nonnullity increased. RT tests evidenced somewhat less power conservatism than that shown by PS, with large power loss only for the 2×2 design if both other effects were nonnull. Note that the slight power advantage of H and M tests is probably due to liberal α .

For $n = 10$ and empirical α results (see Table 3), the M and H tests were slightly liberal for only the main effect test in the 4×4 design. Other tests showed results similar to those found for $n = 5$. Power results were similar

TABLE 2

Normal distribution, n = 5

Stat	Case	Main effect				Case	Interaction		
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4		2 × 2	2 × 4	4 × 4
Empirical Alpha									
F	Null	.0565+	.0528	.0518	.0494+	Null	.0556	.0524	.0506
RO	Null	.0584+	.0516	.0534	.0499+	Null	.0602	.0526	.0518
H	Null	.0589+	.0648	.0602	.0756+	Null	.0604	.0616	.0588
M	Null	.0717+	.0694	.0664	.0784+	Null	.0732	.0674	.0630
RT	1	.0551+	.0506	.0536	.0500+	1	.0570	.0538	.0508
PS	1	.0553+	.0488	.0480	.0472+	1	.0588	.0460	.0420
RT	2	.0531+	.0498	.0524	.0505+	4	.0520	.0566	.0508
PS	2	.0181+	.0218	.0124	.0199+	4	.0184	.0136	.0150
RT	3	.0527+	.0512	.0502	.0506+	7	.0546	.0532	.0534
PS	3	.0048+	.0104	.0036	.0089+	7	.0064	.0038	.0064
RT	4&10	.0534+	.0518	.0522	.0487+	10	.0568	.0542	.0556
PS	4&10	.0177+	.0232	.0184	.0265+	10	.0198	.0196	.0170
RT	5&11	.0211+	.0522	.0450	.0490+	13	.0250	.0538	.0490
PS	5&11	.0056+	.0088	.0050	.0123+	13	.0076	.0058	.0046
RT	6&12	.0117+	.0532	.0428	.0492+	16	.0140	.0516	.0522
PS	6&12	.0019+	.0038	.0016	.0055+	16	.0022	.0020	.0020
RT	7&19	.0534+	.0490	.0520	.0480+	19	.0558	.0530	.0554
PS	7&19	.0063+	.0104	.0070	.0164+	19	.0056	.0090	.0076

TABLE 2 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0130+	.0492	.0444	.0485+	22	.0148	.0520	.0504
PS	8&20	.0016+	.0038	.0026	.0081+	22	.0018	.0016	.0022
RT	9&21	.0043+	.0474	.0404	.0490+	25	.0040	.0518	.0512
PS	9&21	.0002+	.0026	.0010	.0044+	25	.0002	.0006	.0006
<u>Empirical power (medium)</u>									
F	Med.	.8513+	.8488	.8618	.8525+	Med.	.8506	.8500	.8412
RO	Med.	.8442+	.8340	.8450	.8363+	Med.	.8478	.8442	.8340
H	Med.	.8448+	.8616	.8598	.8820+	Med.	.8488	.8594	.8444
M	Med.	.8656+	.8690	.8688	.8871+	Med.	.8694	.8678	.8554
RT	10&4	.8366+	.8332	.8432	.8357+	2	.8380	.8404	.8292
PS	10&4	.8418+	.8344	.8284	.8300+	2	.8426	.8290	.8078
RT	11&5	.8221+	.8254	.8310	.8282+	5	.8222	.8304	.8234
PS	11&5	.6781+	.7242	.6592	.7203+	5	.6716	.6584	.6762
RT	12&6	.8168+	.8240	.8260	.8272+	8	.8140	.8242	.8220
PS	12&6	.5026+	.6190	.4900	.6180+	8	.4958	.4972	.5686
RT	13	.8210+	.8272	.8346	.8336+	11	.8208	.8362	.8220
PS	13	.6764+	.7212	.7206	.7630+	11	.6744	.7176	.6766
RT	14	.6687+	.8076	.8154	.8277+	14	.6606	.8194	.8180
PS	14	.4373+	.6114	.5408	.6432+	14	.4288	.5414	.5428

TABLE 2 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.5464+	.7968	.8076	.8243+	17	.5450	.8128	.8170
PS	15	.2564+	.5046	.3858	.5371+	17	.2510	.3960	.4372
RT	16&22	.8149+	.8276	.8314	.8310+	20	.8144	.8324	.8196
PS	16&22	.5012+	.6226	.5832	.6957+	20	.4990	.5792	.5550
RT	17&23	.5556+	.8070	.8024	.8240+	23	.5436	.8106	.8168
PS	17&23	.2540+	.5094	.4226	.5763+	23	.2516	.4128	.4398
RT	18&24	.3480+	.7932	.7916	.8202+	26	.3398	.8008	.8116
PS	18&24	.1258+	.4112	.2922	.4740+	26	.1202	.2908	.3458
<u>Empirical power (large)</u>									
F	Lar.	.9893+	.9884	.9896	.9875+	Lar.	.9914	.9878	.9900
RO	Lar.	.9860+	.9854	.9884	.9836+	Lar.	.9878	.9868	.9894
H	Lar.	.9862+	.9888	.9904	.9904+	Lar.	.9880	.9892	.9906
M	Lar.	.9897+	.9892	.9906	.9910+	Lar.	.9912	.9902	.9914
RT	19&7	.9864+	.9856	.9870	.9835+	3	.9878	.9862	.9892
PS	19&7	.9867+	.9860	.9846	.9834+	3	.9886	.9838	.9864
RT	20&8	.9827+	.9842	.9870	.9825+	6	.9854	.9838	.9878
PS	20&8	.9500+	.9628	.9436	.9577+	6	.9528	.9458	.9646
RT	21&9	.9810+	.9832	.9876	.9824+	9	.9842	.9830	.9880
PS	21&9	.8823+	.9350	.8786	.9271+	9	.8738	.8852	.9310

TABLE 2 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	22&16	.9827+	.9840	.9856	.9827+	12	.9840	.9840	.9884
PS	22&16	.9495+	.9700	.9626	.9687+	12	.9570	.9598	.9644
RT	23&17	.9437+	.9828	.9856	.9811+	15	.9428	.9802	.9870
PS	23&17	.8333+	.9310	.9070	.9364+	15	.8284	.9028	.9236
RT	24&18	.8895+	.9810	.9830	.9805+	18	.8894	.9792	.9858
PS	24&18	.6738+	.8934	.8222	.8924+	18	.6700	.8244	.8716
RT	25	.9822+	.9850	.9850	.9817+	21	.9830	.9834	.9868
PS	25	.8825+	.9406	.9222	.9525+	21	.8824	.9166	.9314
RT	26	.8931+	.9816	.9818	.9800+	24	.8886	.9784	.9856
PS	26	.6737+	.8978	.8386	.9108+	24	.6666	.8326	.8722
RT	27	.7548+	.9796	.9790	.9797+	27	.7488	.9754	.9846
PS	27	.4787+	.8460	.7408	.8599+	27	.4670	.7400	.8082

+ Entry is averaged over two main effects.

TABLE 3
Normal distribution, n = 10

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
Empirical α									
F	Null	.0500+	.0490	.0492	.0506+	Null	.0540	.0474	.0490
RO	Null	.0519+	.0498	.0484	.0500+	Null	.0554	.0462	.0482
H	Null	.0519+	.0562	.0510	.0616+	Null	.0554	.0508	.0524
M	Null	.0552+	.0574	.0536	.0633+	Null	.0592	.0546	.0544
RT	1	.0508+	.0524	.0510	.0515+	1	.0540	.0490	.0492
PS	1	.0495+	.0516	.0480	.0497+	1	.0528	.0448	.0456
RT	2	.0506+	.0512	.0514	.0523+	4	.0552	.0488	.0504
PS	2	.0295+	.0354	.0298	.0304+	4	.0326	.0280	.0326
RT	3	.0513+	.0506	.0526	.0520+	7	.0570	.0506	.0476
PS	3	.0166+	.0258	.0188	.0216+	7	.0162	.0174	.0216
RT	4&10	.0504+	.0502	.0512	.0513+	10	.0520	.0496	.0518
PS	4&10	.0268+	.0324	.0346	.0358+	10	.0288	.0328	.0302
RT	5&11	.0412+	.0480	.0492	.0503+	13	.0412	.0486	.0522
PS	5&11	.0156+	.0230	.0210	.0232+	13	.0162	.0182	.0220
RT	6&12	.0364+	.0476	.0474	.0511+	16	.0370	.0474	.0524
PS	6&12	.0071+	.0174	.0126	.0174+	16	.0096	.0120	.0148
RT	7&19	.0479+	.0504	.0502	.0505+	19	.0520	.0486	.0514
PS	7&19	.0153+	.0242	.0164	.0273+	19	.0158	.0144	.0206

TABLE 3 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0348+	.0458	.0466	.0507+	22	.0382	.0466	.0530
PS	8&20	.0080+	.0180	.0092	.0176+	22	.0092	.0088	.0134
RT	9&21	.0266+	.0470	.0454	.0508+	25	.0272	.0466	.0522
PS	9&21	.0034+	.0104	.0052	.0138+	25	.0036	.0048	.0084
<u>Empirical power (medium)</u>									
F	Med.	.8514+	.8432	.8488	.8238+	Med.	.8480	.8432	.8608
RO	Med.	.8364+	.8324	.8326	.8094+	Med.	.8354	.8294	.8494
H	Med.	.8364+	.8474	.8428	.8318+	Med.	.8354	.8362	.8568
M	Med.	.8440+	.8498	.8496	.8342+	Med.	.8424	.8414	.8606
RT	10&4	.8339+	.8284	.8344	.8061+	2	.8314	.8272	.8468
PS	10&4	.8283+	.8294	.8272	.8035+	2	.8294	.8200	.8404
RT	11&5	.8309+	.8286	.8312	.8045+	5	.8258	.8236	.8474
PS	11&5	.7624+	.7898	.7556	.7505+	5	.7576	.7544	.7988
RT	12&6	.8266+	.8276	.8290	.8037+	8	.8228	.8244	.8444
PS	12&6	.6816+	.7456	.6824	.7019+	8	.6776	.6858	.7490
RT	13	.8250+	.8282	.8302	.8136+	11	.8270	.8264	.8432
PS	13	.7605+	.7906	.7750	.7684+	11	.7598	.7764	.7942
RT	14	.7999+	.8258	.8268	.7911+	14	.7938	.8180	.8386
PS	14	.6555+	.7392	.6994	.7045+	14	.6558	.6982	.7446

TABLE 3 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.7697+	.8222	.8206	.7873+	17	.7676	.8158	.8382
PS	15	.5508+	.6956	.6234	.6523+	17	.5368	.6270	.6878
RT	16&22	.8210+	.8254	.8284	.8010+	20	.8250	.8228	.8428
PS	16&22	.6802+	.7480	.6696	.7350+	20	.6802	.6714	.7434
RT	17&23	.7659+	.8218	.8188	.7863+	23	.7702	.8136	.8336
PS	17&23	.5455+	.6976	.5890	.6634+	23	.5344	.5886	.6872
RT	18&24	.7074+	.8162	.8144	.7796+	26	.7090	.8090	.8312
PS	18&24	.4182+	.6504	.5170	.6046+	26	.4064	.5162	.6242
<u>Empirical power (large)</u>									
F	Lar.	.9898+	.9994	.9880	.9877+	Lar.	.9896	.9916	.9928
RO	Lar.	.9874+	.9992	.9854	.9858+	Lar.	.9856	.9884	.9908
H	Lar.	.9874+	.9996	.9870	.9887+	Lar.	.9856	.9890	.9910
M	Lar.	.9878+	.9996	.9876	.9890+	Lar.	.9864	.9896	.9912
RT	19&7	.9857+	.9992	.9850	.9851+	3	.9856	.9866	.9906
PS	19&7	.9855+	.9992	.9850	.9848+	3	.9854	.9856	.9894
RT	20&8	.9853+	.9990	.9854	.9838+	6	.9838	.9862	.9894
PS	20&8	.9743+	.9980	.9698	.9746+	6	.9734	.9728	.9840
RT	21&9	.9850+	.9992	.9844	.9831+	9	.9828	.9876	.9896
PS	21&9	.9535+	.9968	.9542	.9645+	9	.9496	.9556	.9750

TABLE 3 (continued)

Stat	Case	Main effect				Case	Interaction		
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4		2 × 2	2 × 4	4 × 4
RT	22&16	.9859+	.9988	.9838	.9848+	12	.9854	.9862	.9904
PS	22&16	.9730+	.9984	.9742	.9791+	12	.9724	.9766	.9844
RT	23&17	.9800+	.9988	.9834	.9825+	15	.9772	.9858	.9896
PS	23&17	.9474+	.9970	.9570	.9656+	15	.9394	.9612	.9732
RT	24&18	.9746+	.9984	.9828	.9807+	18	.9708	.9848	.9882
PS	24&18	.9012+	.9942	.9348	.9530+	18	.8980	.9406	.9628
RT	25	.9845+	.9990	.9830	.9851+	21	.9842	.9860	.9902
PS	25	.9529+	.9966	.9520	.9729+	21	.9488	.9540	.9748
RT	26	.9746+	.9988	.9826	.9814+	24	.9736	.9844	.9886
PS	26	.9016+	.9952	.9254	.9555+	24	.9006	.9248	.9616
RT	27	.9597+	.9988	.9812	.9796+	27	.9548	.9838	.9874
PS	27	.8175+	.9928	.8948	.9405+	27	.8258	.8978	.9464

+Entry is averaged over two main effects.

to those found for $n = 5$, except that the power loss of PS and RT tests generally was not as severe. The empirical α values for the rank tests might differ from .05 because all the rank tests in this study are large-sample tests and only asymptotically distributed as χ^2 or F . For the normal distribution (see Tables 2 and 3), the empirical α values for most rank tests are closer to .05 for $N = 10$ than for $N = 5$.

Exponential Distribution

For $n = 5$ and empirical α results (see Table 4), the M tests were liberal or slightly liberal except for the $df = 3$ main effect in the 2×4 design (empirical $\alpha = .0594$). H tests were slightly liberal or liberal except for main effect and interaction tests in the 2×2 design, the $df = 3$ main effect test in the 2×4 design, and the interaction test in the 4×4 design. F tests were not liberal or conservative, and RO tests were slightly liberal only for tests in the 4×4 design. PS tests showed the same pattern that occurred for the normal distribution: conservatism in the presence of other nonnull effects that worsened as the number of nonnull effects or degree of nonnullity increased. RT tests usually showed the same pattern that occurred for the normal distribution: conservatism in the 2×2 design only if both other effects were nonnull. Exceptions to this conservatism were slightly liberal interaction tests for the 4×4 design, if either main effect was large, and for the 2×4 design, if both main effects were large. Some power results were similar to those for the normal distribution: PS tests with conservatism in the presence of other nonnull effects, RT tests with somewhat less power conservatism than that shown by PS tests, and large power loss only for the 2×2 design if both other effects were nonnull. Power values for F tests are lower than those of RO tests by about 5%–10% in those cases where RO tests had acceptable α . Note that at least some of the power advantage of H and M tests is probably due to slightly liberal or liberal α .

For $n = 10$ and empirical α results (see Table 5), RO tests were not liberal, and the M and H tests were slightly liberal for only the main effect test in the 4×4 design. Other tests showed results similar to those found for $n = 5$. Power results were similar to those found for $n = 5$, except that the power loss of PS and RT tests generally was not as severe and that the power decrement of F , compared to RO, ranged as high as 13%.

Mixed Normal Distribution

For $n = 5$ and empirical α results (see Table 6), F was always conservative. H and M gave acceptable α only for the interaction test in the 2×4 design. Conservatism was evidenced for H and M in main effect and interaction tests for the 2×2 design, the $df = 3$ test in the 2×4 design, and the interaction test in the 4×4 design. However, the $df = 1$ main effect test in the 2×4 design and main effect test in the 4×4 design were liberal for H and M.

TABLE 4
Exponential distribution, n = 5

Stat	Case	2×2	Main effect			Case	Interaction		
			2×4 $df = 1$	2×4 $df = 3$	4×4		2×2	2×4	4×4
<u>Empirical α</u>									
F	Null	.0453+	.0490	.0452	.0466+	Null	.0478	.0472	.0494
RO	Null	.0550+	.0550	.0518	.0648+	Null	.0594	.0582	.0604
H	Null	.0540+	.0660	.0552	.0855+	Null	.0588	.0606	.0590
M	Null	.0644+	.0686	.0594	.0890+	Null	.0690	.0644	.0624
RT	1	.0524+	.0496	.0518	.0519+	1	.0546	.0522	.0532
PS	1	.0520+	.0482	.0448	.0488+	1	.0568	.0440	.0406
RT	2	.0488+	.0518	.0518	.0515+	4	.0536	.0554	.0556
PS	2	.0107+	.0168	.0062	.0102+	4	.0134	.0090	.0096
RT	3	.0468+	.0516	.0492	.0527+	7	.0504	.0562	.0574
PS	3	.0041+	.0068	.0026	.0043+	7	.0042	.0030	.0028
RT	4&10	.0499+	.0524	.0482	.0535+	10	.0542	.0526	.0564
PS	4&10	.0105+	.0134	.0094	.0163+	10	.0128	.0120	.0080
RT	5&11	.0125+	.0538	.0426	.0510+	13	.0138	.0522	.0590
PS	5&11	.0029+	.0060	.0028	.0047+	13	.0038	.0046	.0032
RT	6&12	.0101+	.0512	.0428	.0503+	16	.0112	.0538	.0604
PS	6&12	.0012+	.0024	.0010	.0025+	16	.0024	.0012	.0012
RT	7&19	.0472+	.0514	.0480	.0518+	19	.0512	.0504	.0570
PS	7&19	.0032+	.0036	.0026	.0078+	19	.0044	.0032	.0032

TABLE 4 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0092+	.0512	.0394	.0512+	22	.0104	.0518	.0600
PS	8&20	.0010+	.0024	.0014	.0027+	22	.0014	.0014	.0016
RT	9&21	.0034+	.0518	.0374	.0502+	25	.0036	.0610	.0656
PS	9&21	.0003+	.0012	.0002	.0014+	25	.0004	.0006	.0008
<u>Empirical power (medium)</u>									
F	Med.	.8491+	.8524	.8592	.8450+	Med.	.8530	.8486	.8370
RO	Med.	.9031+	.9328	.9388	.9542+	Med.	.9002	.9436	.9662
H	Med.	.9030+	.9472	.9426	.9676+	Med.	.8998	.9474	.9670
M	Med.	.9166+	.9488	.9462	.9688+	Med.	.9116	.9522	.9698
RT	10&4	.9116+	.9532	.9556	.9812+	2	.9072	.9508	.9704
PS	10&4	.9150+	.9546	.9528	.9800+	2	.9126	.9462	.9650
RT	11&5	.8789+	.9184	.9262	.9534+	5	.8766	.9322	.9610
PS	11&5	.7513+	.8328	.7812	.8718+	5	.7446	.7964	.8746
RT	12&6	.8490+	.8972	.9094	.9366+	8	.8482	.9144	.9472
PS	12&6	.5156+	.6952	.5724	.7420+	8	.5076	.5934	.7472
RT	13	.8762+	.9298	.9350	.9603+	11	.8802	.9378	.9610
PS	13	.7527+	.8486	.8572	.9231+	11	.7600	.8644	.8744
RT	14	.6597+	.8904	.9144	.9396+	14	.6576	.9166	.9442
PS	14	.4210+	.6978	.6622	.7939+	14	.4148	.6622	.7244

TABLE 4 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.5045+	.8710	.8968	.9213+	17	.4994	.9058	.9342
PS	15	.1716+	.5600	.4702	.6375+	17	.1608	.5046	.5830
RT	16&22	.8488+	.9120	.9180	.9456+	20	.8528	.9044	.9440
PS	16&22	.5055+	.7246	.6922	.8465+	20	.5000	.6600	.7414
RT	17&23	.5087+	.8764	.9016	.9298+	23	.4948	.8846	.9348
PS	17&23	.1759+	.5818	.5274	.7045+	23	.1608	.4724	.5844
RT	18&24	.1882+	.8622	.8798	.9108+	26	.1774	.8760	.9282
PS	18&24	.0424+	.4620	.3578	.5548+	26	.0360	.3328	.4550
<u>Empirical power (large)</u>									
F	Lar.	.9711+	.9782	.9742	.9794+	Lar.	.9664	.9738	.9806
RO	Lar.	.9847+	.9956	.9944	.9977+	Lar.	.9818	.9936	.9990
H	Lar.	.9846+	.9956	.9948	.9988+	Lar.	.9818	.9942	.9992
M	Lar.	.9877+	.9962	.9954	.9988+	Lar.	.9848	.9954	.9992
RT	19&7	.9866+	.9972	.9956	.9994+	3	.9834	.9948	.9988
PS	19&7	.9876+	.9966	.9946	.9993+	3	.9842	.9944	.9988
RT	20&8	.9856+	.9950	.9932	.9973+	6	.9850	.9952	.9990
PS	20&8	.9706+	.9872	.9756	.9930+	6	.9668	.9782	.9946
RT	21&9	.9767+	.9906	.9902	.9963+	9	.9732	.9924	.9984
PS	21&9	.8988+	.9626	.9240	.9740+	9	.8892	.9412	.9834

TABLE 4 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	22&16	.9857+	.9952	.9940	.9982+	12	.9828	.9952	.9990
PS	22&16	.9697+	.9896	.9848	.9958+	12	.9666	.9854	.9952
RT	23&17	.9520+	.9906	.9894	.9968+	15	.9452	.9928	.9984
PS	23&17	.8867+	.9662	.9432	.9831+	15	.8826	.9502	.9818
RT	24&18	.9123+	.9856	.9886	.9957+	18	.9076	.9900	.9972
PS	24&18	.7318+	.9238	.8882	.9558+	18	.7160	.9000	.9530
RT	25	.9761+	.9922	.9916	.9971+	21	.9732	.9926	.9984
PS	25	.8959+	.9696	.9582	.9896+	21	.8956	.9628	.9826
RT	26	.9138+	.9884	.9884	.9953+	24	.9080	.9886	.9976
PS	26	.7261+	.9406	.8974	.9672+	24	.7228	.9010	.9528
RT	27	.7377+	.9836	.9874	.9947+	27	.7302	.9878	.9964
PS	27	.4591+	.8814	.8210	.9323+	27	.4518	.8284	.8994

+ Entry is averaged over two main effects

TABLE 5
Exponential distribution, n = 10

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
<u>Empirical α</u>									
F	Null	.0440+	.0456	.0434	.0504+	Null	.0518	.0506	.0444
RO	Null	.0478+	.0466	.0514	.0572+	Null	.0570	.0562	.0504
H	Null	.0463+	.0498	.0512	.0655+	Null	.0558	.0558	.0482
M	Null	.0507+	.0508	.0538	.0664+	Null	.0580	.0594	.0500
RT	1	.0465+	.0444	.0480	.0524+	1	.0578	.0494	.0434
PS	1	.0450+	.0430	.0454	.0508+	1	.0556	.0472	.0394
RT	2	.0482+	.0470	.0496	.0525+	4	.0576	.0542	.0474
PS	2	.0194+	.0220	.0144	.0227+	4	.0202	.0196	.0222
RT	3	.0461+	.0480	.0486	.0505+	7	.0548	.0532	.0488
PS	3	.0102+	.0150	.0072	.0132+	7	.0096	.0128	.0104
RT	4&10	.0469+	.0486	.0490	.0521+	10	.0552	.0536	.0502
PS	4&10	.0173+	.0238	.0218	.0307+	10	.0210	.0228	.0218
RT	5&11	.0330+	.0494	.0480	.0501+	13	.0384	.0558	.0534
PS	5&11	.0086+	.0132	.0078	.0150+	13	.0098	.0120	.0130
RT	6&12	.0286+	.0494	.0488	.0491+	16	.0302	.0592	.0548
PS	6&12	.0041+	.0102	.0042	.0090+	16	.0048	.0076	.0068
RT	7&19	.0465+	.0486	.0498	.0525+	19	.0526	.0510	.0496
PS	7&19	.0086+	.0166	.0050	.0194+	19	.0092	.0074	.0122

TABLE 5 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0285+	.0468	.0460	.0506+	22	.0308	.0518	.0562
PS	8&20	.0044+	.0094	.0028	.0102+	22	.0040	.0038	.0068
RT	9&21	.0200+	.0476	.0450	.0502+	25	.0178	.0612	.0588
PS	9&21	.0028+	.0080	.0022	.0063+	25	.0024	.0032	.0036
Empirical power (medium)									
F	Med.	.8482+	.8554	.8544	.8223+	Med.	.8428	.8464	.8528
RO	Med.	.9486+	.9652	.9698	.9698+	Med.	.9394	.9694	.9878
H	Med.	.9479+	.9684	.9716	.9741+	Med.	.9390	.9698	.9878
M	Med.	.9512+	.9692	.9730	.9748+	Med.	.9424	.9712	.9880
RT	10	.9570+	.9792	.9798	.9869+	2	.9486	.9742	.9902
PS	10	.9551+	.9802	.9796	.9868+	2	.9474	.9732	.9898
RT	11	.9303+	.9568	.9596	.9687+	5	.9226	.9612	.9834
PS	11	.8814+	.9364	.9164	.9413+	5	.8758	.9212	.9662
RT	12	.9042+	.9444	.9464	.9551+	8	.9008	.9466	.9756
PS	12	.7666+	.8798	.8284	.8824+	8	.7592	.8418	.9328
RT	13	.9302+	.9620	.9652	.9758+	11	.9206	.9640	.9846
PS	13	.8791+	.9396	.9410	.9622+	11	.8678	.9378	.9650
RT	14	.8257+	.9366	.9540	.9591+	14	.8176	.9452	.9750
PS	14	.6955+	.8806	.8756	.9064+	14	.6836	.8668	.9326

TABLE 5 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.7545+	.9210	.9406	.9444+	17	.7472	.9384	.9688
PS	15	.4809+	.8160	.7744	.8390+	17	.4738	.7992	.8824
RT	16	.9037+	.9468	.9460	.9644+	20	.8942	.9378	.9764
PS	16	.7629+	.8916	.8180	.9298+	20	.7582	.8110	.9294
RT	17	.7499+	.9270	.9368	.9495+	23	.7396	.9232	.9688
PS	17	.4847+	.8256	.7360	.8609+	23	.4808	.7184	.8810
RT	18	.5818+	.9144	.9216	.9352+	26	.5818	.9154	.9628
PS	18	.2382+	.7600	.6366	.7896+	26	.2418	.6416	.8256
<u>Empirical power (large)</u>									
F	Lar.	.9766+	.9974	.9816	.9834+	Lar.	.9740	.9846	.9858
RO	Lar.	.9961+	1.0000	.9994	.9995+	Lar.	.9946	.9992	.9998
H	Lar.	.9959+	1.0000	.9994	.9997+	Lar.	.9946	.9992	.9998
M	Lar.	.9964+	1.0000	.9994	.9997+	Lar.	.9950	.9992	.9998
RT	19	.9963+	1.0000	.9996	.9999+	3	.9960	.9990	.9996
PS	19	.9965+	1.0000	.9996	.9999+	3	.9956	.9990	.9996
RT	20	.9961+	1.0000	.9990	.9996+	6	.9952	.9982	.9998
PS	20	.9925+	1.0000	.9966	.9985+	6	.9910	.9972	.9994
RT	21	.9933+	1.0000	.9976	.9992+	9	.9920	.9978	.9998
PS	21	.9785+	.9998	.9882	.9964+	9	.9734	.9922	.9988

TABLE 5 (continued)

Stat	Case	Main effect				Case	Interaction		
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4		2 × 2	2 × 4	4 × 4
RT	22	.9950+	1.0000	.9990	.9994+	12	.9952	.9988	.9998
PS	22	.9914+	1.0000	.9978	.9988+	12	.9906	.9978	.9994
RT	23	.9869+	1.0000	.9972	.9992+	15	.9842	.9980	.9996
PS	23	.9709+	.9992	.9920	.9965+	15	.9684	.9928	.9984
RT	24	.9719+	.9998	.9964	.9980+	18	.9698	.9966	.9996
PS	24	.9224+	.9980	.9822	.9940+	18	.9200	.9848	.9968
RT	25	.9927+	1.0000	.9974	.9986+	21	.9924	.9980	.9996
PS	25	.9773+	.9996	.9880	.9976+	21	.9748	.9872	.9988
RT	26	.9730+	.9998	.9962	.9979+	24	.9712	.9954	.9996
PS	26	.9239+	.9980	.9766	.9946+	24	.9198	.9726	.9962
RT	27	.9326+	1.0000	.9964	.9974+	27	.9260	.9942	.9994
PS	27	.8056+	.9968	.9630	.9903+	27	.8022	.9536	.9906

+Entry is averaged over two main effects.

TABLE 6
Mixed normal distribution, n = 5

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
<u>Empirical α</u>									
F	Null	.0237+	.0188	.0226	.0291+	Null	.0192	.0186	.0410
RO	Null	.0314+	.2486	.0320	.4514+	Null	.0282	.1212	.1372
H	Null	.0293+	.2172	.0200	.3537+	Null	.0254	.0434	.0208
M	Null	.0345+	.2244	.0228	.3595+	Null	.0330	.0482	.0220
RT	1	.0503+	.0532	.0534	.0522+	1	.0458	.0486	.0538
PS	1	.0516+	.0494	.0448	.0496+	1	.0472	.0438	.0404
RT	2	.0461+	.0482	.0510	.0493+	4	.0410	.0360	.0534
PS	2	.0019+	.0004	.0002	.0000+	4	.0018	.0002	.0000
RT	3	.0462+	.0482	.0506	.0497+	7	.0414	.0328	.0504
PS	3	.0015+	.0004	.0002	.0000+	7	.0014	.0000	.0000
RT	4&10	.0480+	.0410	.0480	.0442+	10	.0412	.0438	.0572
PS	4&10	.0025+	.0002	.0004	.0000+	10	.0018	.0010	.0002
RT	5&11	.0024+	.0290	.0316	.0336+	13	.0024	.2122	.1512
PS	5&11	.0013+	.0006	.0006	.0000+	13	.0010	.0004	.0008
RT	6&12	.0037+	.0100	.0240	.0289+	16	.0036	.1902	.1662
PS	6&12	.0012+	.0006	.0006	.0000+	16	.0012	.0004	.0006
RT	7&19	.0480+	.0392	.0480	.0397+	19	.0412	.0436	.0546
PS	7&19	.0022+	.0002	.0004	.0000+	19	.0012	.0010	.0002

TABLE 6 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0030+	.0130	.0316	.0351+	22	.0034	.1616	.1704
PS	8&20	.0012+	.0004	.0006	.0000+	22	.0010	.0002	.0004
RT	9&21	.0030+	.0220	.0328	.0299+	25	.0030	.3620	.2450
PS	9&21	.0010+	.0004	.0006	.0000+	25	.0010	.0004	.0004
<u>Empirical power (medium)</u>									
F	Med.	.8037+	.8184	.8212	.8269+	Med.	.8040	.8244	.8212
RO	Med.	.9780+	.9924	.9978	.9999+	Med.	.9814	.9984	1.0000
H	Med.	.9772+	.9928	.9976	.9998+	Med.	.9800	.9984	1.0000
M	Med.	.9789+	.9930	.9976	.9998+	Med.	.9828	.9984	1.0000
RT	10&4	.9891+	1.0000	1.0000	1.0000+	2	.9904	1.0000	1.0000
PS	10&4	.9885+	.9998	1.0000	1.0000+	2	.9906	1.0000	1.0000
RT	11&5	.9434+	.9904	.9982	1.0000+	5	.9466	.9988	1.0000
PS	11&5	.8814+	.9644	.9780	.9994+	5	.8876	.9904	1.0000
RT	12&6	.8404+	.9806	.9874	.9995+	8	.8420	.9936	1.0000
PS	12&6	.2212+	.8672	.8234	.9901+	8	.2176	.8932	.9956
RT	13	.9436+	.9986	.9990	1.0000+	11	.9454	.9994	1.0000
PS	13	.8834+	.9926	.9954	.9999+	11	.8820	.9970	1.0000
RT	14	.6434+	.9842	.9948	.9999+	14	.6464	.9952	1.0000
PS	14	.4109+	.9162	.9366	.9950+	14	.4150	.9574	.9954

TABLE 6 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.4561+	.9754	.9858	.9979+	17	.4592	.9946	1.0000
PS	15	.0141+	.8384	.7392	.9175+	17	.0182	.9426	.9616
RT	16&22	.8397+	.9920	.9930	1.0000+	20	.8422	.9728	1.0000
PS	16&22	.2227+	.9498	.9304	.9990+	20	.2184	.7620	.9974
RT	17&23	.4519+	.9738	.9906	.9993+	23	.4576	.9788	1.0000
PS	17&23	.0153+	.8322	.8664	.9707+	23	.0140	.6876	.9626
RT	18&24	.0015+	.9626	.9766	.9940+	26	.0022	.9778	.9990
PS	18&24	.0002+	.7152	.5982	.7699+	26	.0002	.6348	.9096
<u>Empirical power (large)</u>									
F	Lar.	.9282+	.9518	.9454	.9556+	Lar.	.9308	.9462	.9556
RO	Lar.	.9916+	.9988	.9996	1.0000+	Lar.	.9924	.9998	1.0000
H	Lar.	.9915+	.9990	.9996	1.0000+	Lar.	.9922	.9996	1.0000
M	Lar.	.9917+	.9990	.9996	1.0000+	Lar.	.9926	.9998	1.0000
RT	19&7	.9912+	1.0000	1.0000	1.0000+	3	.9926	1.0000	1.0000
PS	19&7	.9906+	1.0000	1.0000	1.0000+	3	.9926	1.0000	1.0000
RT	20&8	.9936+	.9998	.9998	1.0000+	6	.9952	1.0000	1.0000
PS	20&8	.9918+	.9996	.9994	1.0000+	6	.9936	.9992	1.0000
RT	21&9	.9627+	.9926	.9990	1.0000+	9	.9624	.9996	1.0000
PS	21&9	.9039+	.9674	.9850	1.0000+	9	.9086	.9948	1.0000

TABLE 6 (continued)

Stat	Case	Main effect				Case	Interaction		
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4		2 × 2	2 × 4	4 × 4
RT	22&16	.9935+	.9996	.9998	1.0000+	12	.9948	1.0000	1.0000
PS	22&16	.9913+	.9996	.9996	1.0000+	12	.9936	.9998	1.0000
RT	23&17	.9931+	.9990	.9970	1.0000+	15	.9948	.9994	1.0000
PS	23&17	.9920+	.9962	.9734	1.0000+	15	.9948	.9956	1.0000
RT	24&18	.9609+	.9930	.9976	1.0000+	18	.9632	.9982	1.0000
PS	24&18	.9015+	.9614	.9714	.9991+	18	.9040	.9838	1.0000
RT	25	.9627+	.9988	.9996	1.0000+	21	.9620	.9996	1.0000
PS	25	.9049+	.9974	.9968	1.0000+	21	.9050	.9984	1.0000
RT	26	.9600+	.9990	.9970	1.0000+	24	.9616	.9986	1.0000
PS	26	.8991+	.9946	.9606	.9997+	24	.9004	.9870	1.0000
RT	27	.6796+	.9914	.9974	1.0000+	27	.6822	.9976	1.0000
PS	27	.4328+	.9460	.9574	.9972+	27	.4344	.9674	.9956

+ Entry is averaged over two main effects.

RO evidenced a similar pattern to H and M, except that the liberal nature of the tests extended to interaction tests in the 2×4 and 4×4 designs. PS showed the same pattern that occurred for the normal distribution: conservatism in the presence of other nonnull effects which worsened as the number of nonnull effects or degree of nonnullity increased. RT usually showed conservatism (which was worse for the 2×2 design) only if both other effects were nonnull, except for liberal interaction tests for the 2×4 and 4×4 designs if both main effects were nonnull. Power results show F to be conservative by as much as 18% for medium power for the 2×2 design as compared to RO. However, RO, H, M, and RT all have some problem with liberal α , invalidating many power comparisons for all but the 2×2 design. Additionally, in the 2×2 design, RT suffers from a power decrement in the presence of both other effects being nonnull. PS is conservative in the presence of other nonnull effects.

For $n = 10$ and empirical α (see Table 7), results are very similar to those for $n = 5$. The liberal patterns for RO, H, and M are not as severe, and α for F seems to become less conservative as the size of the design increases. Power results mirror those for $n = 5$.

Conclusions

Researchers desiring an alternative to the parametric fixed effects two-way ANOVA will find little consolation from the results of this simulation study. While the PS, H, and M tests all share admirable properties in the theory of their development, they also share, along with RT, some serious problems. PS shows conservatism of both α and power in the presence of other nonnull effects which worsen as the number of nonnull effects or degree of nonnullity increases, consistent with the dampening effect found by Toothaker and Chang (1980) and Lemmer (1980). This lack of independence of the tests makes PS a poor choice. H has liberal α for some designs and distributions, notably for two cases: small n with normal or exponential distributions and the larger designs when samples are taken from a mixed normal distribution. M shares many of the problems of H, with even more liberal results for small n . RO, in spite of its intuitive appeal, also shares many of the problems of H. Although RO is not as liberal for small n from normal or exponential distributions, it is more liberal in some designs when samples were taken from the mixed normal distribution. RT shares many of the problems of PS, although not to the same degree. Even though the rank tests have asymptotic distributions with known formulas for degrees of freedom, perhaps the small-sample distributions of the statistics would be better fit by χ^2 or F distributions with different (adjusted) degrees of freedom. In contrast, F suffers from conservatism with the mixed normal distribution, but it never exhibits a liberal α .

TABLE 7

Mixed normal distribution, n = 10

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
<u>Empirical α</u>									
F	Null	.0177+	.0330	.0294	.0406+	Null	.0156	.0238	.0372
RO	Null	.0351+	.1532	.0394	.2529+	Null	.0368	.0882	.1030
H	Null	.0200+	.1036	.0150	.1470+	Null	.0244	.0268	.0186
M	Null	.0218+	.1056	.0158	.1484+	Null	.0260	.0288	.0186
RT	1	.0504+	.0460	.0478	.0492+	1	.0532	.0504	.0514
PS	1	.0480+	.0472	.0460	.0483+	1	.0516	.0472	.0458
RT	2	.0465+	.0498	.0504	.0518+	4	.0492	.0486	.0500
PS	2	.0018+	.0024	.0004	.0000+	4	.0020	.0002	.0002
RT	3	.0468+	.0496	.0506	.0510+	7	.0494	.0450	.0524
PS	3	.0016+	.0022	.0004	.0000+	7	.0020	.0002	.0000
RT	4&10	.0477+	.0462	.0458	.0503+	10	.0518	.0452	.0518
PS	4&10	.0017+	.0006	.0000	.0007+	10	.0018	.0002	.0004
RT	5&11	.0036+	.0452	.0294	.0499+	13	.0028	.1740	.1638
PS	5&11	.0009+	.0004	.0006	.0000+	13	.0008	.0006	.0004
RT	6&12	.0077+	.0308	.0286	.0486+	16	.0078	.1778	.1838
PS	6&12	.0009+	.0004	.0004	.0000+	16	.0008	.0006	.0004
RT	7&19	.0473+	.0454	.0470	.0495+	19	.0514	.0456	.0498
PS	7&19	.0015+	.0002	.0000	.0000+	19	.0018	.0002	.0002

TABLE 7 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	8&20	.0088+	.0290	.0284	.0482+	22	.0056	.0616	.1840
PS	8&20	.0009+	.0000	.0004	.0000+	22	.0008	.0002	.0006
RT	9&21	.0044+	.0394	.0322	.0469+	25	.0028	.2948	.2888
PS	9&21	.0009+	.0000	.0002	.0000+	25	.0008	.0004	.0006
Empirical power (medium)									
F	Med.	.8228+	.8308	.8228	.8103+	Med.	.8144	.8336	.8362
RO	Med.	.9982+	.9992	1.0000	1.0000+	Med.	.9974	1.0000	1.0000
H	Med.	.9981+	.9990	1.0000	1.0000+	Med.	.9972	1.0000	1.0000
M	Med.	.9982+	.9990	1.0000	1.0000+	Med.	.9974	1.0000	1.0000
RT	10&4	1.0000+	1.0000	1.0000	1.0000+	2	1.0000	1.0000	1.0000
PS	10&4	1.0000+	1.0000	1.0000	1.0000+	2	1.0000	1.0000	1.0000
RT	11&5	.9985+	1.0000	1.0000	1.0000+	5	.9990	1.0000	1.0000
PS	11&5	.9960+	1.0000	.9998	1.0000+	5	.9964	1.0000	1.0000
RT	12&6	.9848+	.9998	.9998	1.0000+	8	.9842	1.0000	1.0000
PS	12&6	.9125+	.9984	.9974	1.0000+	8	.9120	.9998	1.0000
RT	13	.9980+	1.0000	1.0000	1.0000+	11	.9990	1.0000	1.0000
PS	13	.9959+	1.0000	1.0000	1.0000+	11	.9956	1.0000	1.0000
RT	14	.8913+	.9998	1.0000	1.0000+	14	.8910	1.0000	1.0000
PS	14	.7655+	.9990	1.0000	1.0000+	14	.7708	.9998	1.0000

TABLE 7 (continued)

Stat	Case	Main effect				Interaction			
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4	Case	2 × 2	2 × 4	4 × 4
RT	15	.6697+	.9992	1.0000	1.0000+	17	.6718	1.0000	1.0000
PS	15	.1512+	.9910	.9944	.9995+	17	.1564	.9994	1.0000
RT	16&22	.9860+	1.0000	1.0000	1.0000+	20	.9832	.9988	1.0000
PS	16&22	.9121+	.9998	.9974	1.0000+	20	.9096	.9864	1.0000
RT	17&23	.6687+	.9998	1.0000	1.000+	23	.6804	.9990	1.0000
PS	17&23	.1516+	.9932	.9916	.9999+	23	.1610	.9750	1.0000
RT	18&24	.0190+	.9978	.9994	1.0000+	26	.0218	.9990	1.0000
PS	18&24	.0006+	.9732	.9638	.9983+	26	.0006	.9582	1.0000
<u>Empirical power (large)</u>									
F	Lar.	.9515+	.9898	.9624	.9687+	Lar.	.9516	.9642	.9740
RO	Lar.	.9998+	1.0000	1.0000	1.0000+	Lar.	.9998	1.0000	1.0000
H	Lar.	.9998+	1.0000	1.0000	1.0000+	Lar.	.9998	1.0000	1.0000
M	Lar.	.9998+	1.0000	1.0000	1.0000+	Lar.	.9998	1.0000	1.0000
RT	19&7	1.0000+	1.0000	1.0000	1.0000+	3	1.0000	1.0000	1.0000
PS	19&7	1.0000+	1.0000	1.0000	1.0000+	3	1.0000	1.0000	1.0000
RT	20&8	1.0000+	1.0000	1.0000	1.0000+	6	1.0000	1.0000	1.0000
PS	20&8	1.0000+	1.0000	1.0000	1.0000+	6	1.0000	1.0000	1.0000
RT	21&9	.9991+	1.0000	1.0000	1.0000+	9	.9992	1.0000	1.0000
PS	21&9	.9971+	1.0000	1.0000	1.0000+	9	.9972	1.0000	1.0000

TABLE 7 (continued)

Stat	Case	Main effect				Case	Interaction		
		2 × 2	2 × 4 <i>df</i> = 1	2 × 4 <i>df</i> = 3	4 × 4		2 × 2	2 × 4	4 × 4
RT	22&16	1.0000+	1.0000	1.0000	1.0000+	12	1.0000	1.0000	1.0000
PS	22&16	1.0000+	1.0000	1.0000	1.0000+	12	1.0000	1.0000	1.0000
RT	23&17	1.0000+	1.0000	1.0000	1.0000+	15	1.0000	1.0000	1.0000
PS	23&17	1.0000+	1.0000	1.0000	1.0000+	15	1.0000	1.0000	1.0000
RT	24&18	.9981+	1.0000	1.0000	1.0000+	18	.9978	1.0000	1.0000
PS	24&18	.9939+	1.0000	1.0000	1.0000+	18	.9930	1.0000	1.0000
RT	25	.9990+	1.0000	1.0000	1.0000+	21	.9994	1.0000	1.0000
PS	25	.9967+	1.0000	1.0000	1.0000+	21	.9968	.9998	1.0000
RT	26	.9979+	1.0000	1.0000	1.0000+	24	.9970	1.0000	1.0000
PS	26	.9925+	1.0000	1.0000	1.0000+	24	.9938	1.0000	1.0000
RT	27	.9073+	1.0000	1.0000	1.0000+	27	.9082	1.0000	1.0000
PS	27	.7819+	1.0000	1.0000	1.0000+	27	.7888	.9994	1.0000

+Entry is averaged over two main effects.

If a researcher has some knowledge of the type of distribution for the population sampled, a choice may be made from among F , RO, and H, if the cell size is not too small and if the distribution is similar to the normal or exponential distribution. If the distribution is suspected to be like the mixed normal, only the F can be recommended unless the design is 2×2 , in which case RO, H, and M all have adequate control of α and give higher power than F .

Certainly more research is needed for these procedures when the equal variance assumption has been violated, especially in combination with non-normal distributions of the type often encountered in the real world by researchers. Also, for the rank tests, more work is needed in the following areas: using estimates other than the mean in forming appropriate quantities for aligning and using simple main effect analysis that relies on the qualities of ranking in a one-way design. Finally, more research is needed comparing other transformations on the data (e.g., power transformations) to ranking and use of robust estimators in two-way analyses.

References

- Akritas, M. G. (1990). The rank transform method in some two-factor designs. *Journal of the American Statistical Association*, 85, 73–78.
- Blair, R. C., Sawilowsky, S. S., & Higgins, J. J. (1987). Limitations of the rank transform statistic in tests for interactions. *Communications in Statistics: Simulation and Computation*, B16, 1133–1145.
- Box, G. E. P., & Muller, M. E. (1958). A note on the generation of random normal deviates. *Annals of Mathematical Statistics*, 29, 610–611.
- Chen, E. H. (1971). Random normal number generator for 32-bit-word computers. *Journal of the American Statistical Association*, 66, 400–403.
- Conover, W. J., & Iman, R. L. (1976). On some alternative procedures using ranks for the analysis of experimental designs. *Communications in Statistics*, A5, 1349–1368.
- Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35(3), 124–129.
- Fawcett, R. F., & Salter, K. C. (1984). A Monte Carlo study of the F test and three tests based on ranks of treatment effects in randomized block designs. *Communications in Statistics: Simulation and Computation*, 13(2), 213–225.
- Fligner, M. A. (1981). Comment. *American Statistician*, 35(3), 131–132.
- Groggel, D. J. (1987). A Monte Carlo study of rank tests for block designs. *Communications in Statistics: Simulation and Computation*, 16(3), 601–620.
- Harwell, M. R. (1991). Completely randomized factorial analysis of variance using ranks. *British Journal of Mathematical and Statistical Psychology*, 44, 383–401.
- Harwell, M. R., & Serlin, R. C. (1990). A nonparametric test statistic for the general linear model. *Journal of Educational Statistics*, 14, 351–371.
- Hettmansperger, T. P. (1984). *Statistical inference based on ranks*. New York: Wiley.
- Hodges, J. L., Jr., & Lehmann, E. L. (1962). Rank methods for combination of independent experiments in the analysis of variance. *Annals of Mathematical Statistics*, 33, 482–497.

- Hora, S. C., & Conover, W. J. (1984). The F statistic in the two-way layout with rank-score transformed data. *Journal of the American Statistical Association*, 79, 668–673.
- Hora, S. C., & Iman, R. L. (1988). Asymptotic relative efficiencies of the rank transformation procedures in randomized complete block designs. *Journal of the American Statistical Association*, 83, 462–470.
- Iman, R. L. (1974). A power study of a rank transform for the two-way classification model when interaction may be present. *Canadian Journal of Statistics*, 2, 227–239.
- Iman, R. L., & Conover, W. J. (1976). *A comparison of several rank tests for the two-way layout* (SAND76-0631). Albuquerque, NM: Sandia Laboratories.
- Iman, R. L., Hora, S. C., & Conover, W. J. (1984). Comparison of asymptotically distribution-free procedures for the analysis of complete blocks. *Journal of the American Statistical Association*, 79, 674–685.
- Koch, G. G. (1969). Some aspects of the statistical analysis of “split-plot” experiments in completely randomized layouts. *Journal of the American Statistical Association*, 64, 485–505.
- Koch, G. G., & Sen, P. K. (1968). Some aspects of the statistical analysis of the “mixed model.” *Biometrics*, 24, 27–48.
- Lehmann, R. S., & Bailey, D. E. (1968). *Digital computing: Fortran IV and its applications in behavioral science*. New York: Wiley.
- Lemmer, H. H. (1980). Some empirical results on the two-way analysis of variance with ranks. *Communications in Statistics: Theories and Methods*, A9, 1427–1438.
- McSweeney, M. (1967). An empirical study of two proposed nonparametric tests for main effects and interaction (Doctoral dissertation, University of California–Berkeley, 1968). *Dissertation Abstracts International*, 28(10), 4005.
- Mehra, K. L., & Sarangi, J. (1967). Asymptotic efficiency of certain rank tests for comparative experiments. *The Annals of Mathematical Statistics*, 38, 90–107.
- Mehra, K. L., & Sen, P. K. (1969). On a class of conditionally distribution-free tests for interactions in factorial experiments. *The Annals of Mathematical Statistics*, 40, 658–664.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156–166.
- Puri, M. L., & Sen, P. K. (1985). *Nonparametric methods in general linear models*. New York: Wiley.
- Reinach, S. G. (1965). A nonparametric analysis for a multiway classification with one element per cell. *South African Journal of Agricultural Science*, 8, 941–960.
- SAS Institute. (1990). *SAS procedures guide, Version 6* (3rd ed.). Cary, NC: Author.
- Sawilowsky, S. S. (1985). Robust and power analysis of the $2 \times 2 \times 2$ ANOVA, rank transformation, random normal scores, and expected normal scores transformation tests (Doctoral dissertation, University of South Florida, 1986). *Dissertation Abstracts International*, 46(10), 3009A.
- Sawilowsky, S. S. (1989, April). *Rank transform: The bridge is falling down*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Sawilowsky, S. S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research*, 60, 91–126.

- Sawilowsky, S. S., Blair, R. C., & Higgins, J. J. (1989). An investigation of the Type I error and power properties of the rank transformation procedure in factorial ANOVA. *Journal of Educational Statistics, 14*, 255–267.
- Scheirer, C. J., Ray, W. S., & Hare, N. (1976). The analysis of ranked data derived from completely randomized factorial designs. *Biometrics, 32*, 429–434.
- Sen, P. K. (1968). On a class of aligned rank order tests in two-way layouts. *The Annals of Mathematical Statistics, 39*, 1115–1124.
- Thompson, G. L., & Ammann, L. P. (1989). Efficiencies of the rank-transform in two-way models with no interaction. *Journal of the American Statistical Association, 84*(405), 325–330.
- Toothaker, L. E., & Chang, H. (1980). On the analysis of ranked data derived from completely randomized factorial designs. *Journal of Educational Statistics, 5*, 169–176.

Authors

LARRY E. TOOTHAKER is David Ross Boyd Professor of Psychology at the University of Oklahoma, Norman, OK 73019. He specializes in robustness of ANOVA, including repeated measures designs, multiple comparison procedures, and nonparametrics.

DE A. NEWMAN is a PhD Candidate, University of Oklahoma, Norman, OK 73019. She specializes in robustness studies, multiple comparisons, and nonparametrics.

Received December 2, 1992

Revision received September 30, 1993

Accepted December 1, 1993