ADVANCES IN MULTIPLE COMPARISONS
IN THE LAST DECADE

BY

RUPERT G. MILLER, JR.

TECHNICAL REPORT NO. 26
OCTOBER 21, 1976

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 RO1 GM21215-02

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ADVANCES IN MULTIPLE COMPARISONS
IN THE LAST DECADE

BY

RUPERT G. MILLER, JR.

TECHNICAL REPORT NO. 26
OCTOBER 21, 1976

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 RO1 GM21215-02

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ADVANCES IN MULTIPLE COMPARISONS  
IN THE LAST DECADE  

Rupert G. Miller, Jr.+  
Stanford University  

This article contains a bibliography of papers on 
multiple comparisons published between 1966 and 1976. A 
discussion of some of the more important developments 
during this period precedes the bibliography.  

1. Introduction  

was published. It attempted to summarize what was then known about  
the theory and methods of multiple comparisons. During the ensuing  
decade research in this field has continued and a variety of new  
results have been obtained.  

A session on multiple comparisons was planned for the 1976  
Boston ASA meetings. I was asked to participate and give an overview  
of developments in the field in the last ten years. This paper is an  
outgrowth of preparation for that talk.  

I decided the most important contribution I could make would  
be to compile a bibliography of papers written on multiple comparisons  

+Research supported by Public Health Service Grant 5 R01 GM21215-02
since the publication of my book. This was accomplished by scanning most of the statistical journals for the years 1966 through 1975 and those issues for 1976 which were available at the time. In addition, a number of statistical symposia and festschrift and some psychological journals which publish statistical articles were examined as well. A few articles in non-statistical journals were brought to my attention by other statisticians. The papers collected in this search are listed in Section 4.

Some papers were excluded which might have been included under different circumstances. In particular, papers on ranking and selection were omitted. These are involved with multiple comparisons questions, but the general purpose of ranking and selection methods and the probabilities which they control are slightly different from the multiple comparisons techniques considered here. An additional deterrent to including them was the large number of these papers. A separate bibliography is required for ranking and selection.

Also, papers on outlier detection have been excluded. Early techniques in this field employed statistics such as the studentized extreme deviate from the sample mean which are clearly multiple comparisons statistics. However, recent work on outliers has shifted into other directions and more emphasis is placed on automatically handling outliers through robust estimators than in detecting them.

For a few papers it was difficult to decide whether they should be considered to be on multiple comparisons or not. This occurred principally for papers on distribution theory in multivariate
analysis where the statistics involved could be used for multiple comparisons. If the paper was concerned exclusively with analytic distribution theory and did not mention the multiple comparisons problem or include a table for use with the multiple comparisons test, then it was excluded.

2. Papers of Special Interest

The papers in the bibliography have been grouped according to their primary topic. Research has been more active on some of these topics than on others, and some topics are of greater interest to me than others. There are many worthwhile papers in the bibliography, but I would like to discuss briefly those topics and papers which in my opinion have the greatest interest and potential usefulness.

2.1. Probability Inequalities

At the time I wrote my book I knew the Bonferroni inequality \( P(\bigcap_{\mathbb{1}}^{x} A_{\mathbb{1}}) \geq 1 - \sum_{\mathbb{1}}^{x} P(A_{\mathbb{1}}^{C}) \) was very useful, but over the course of the past ten years I have become even more impressed with the tightness of the results it gives if the \( P(A_{\mathbb{1}}^{C}) \) are small. Although special techniques and distribution theory can improve on it, the improvement is very often only minor. This is true whether the Bonferroni inequality is used to simply combine multiple t tests (cf., [6, pp. 67-70]) or to weld together separate confidence intervals on different parameters for a confidence interval on a function of all the parameters as in tolerance intervals about regression lines (cf., [5]).
Nonetheless, Šidák [24] gave the proof for a general inequality which gives slight improvement over the Bonferroni inequality when both are applicable (cf., [48] and [62]) and which is useful in the proofs of some theorems concerned with multiple comparisons techniques (e.g., [49]). The inequality states that if the random vector \( \mathbf{Y} = (Y_1, \ldots, Y_r) \) has a multivariate normal distribution \( N(0, \Sigma) \) with zero mean vector and arbitrary covariance matrix \( \Sigma \) and if \( S^2 \) is an independent random variable distributed as \( \chi^2_\nu / \nu \), then

\[
P \left\{ \frac{|Y_1|}{S} \leq c_1, \ldots, \frac{|Y_r|}{S} \leq c_r \right\} \geq \frac{r}{\nu} \prod_{i=1}^{r} P \left\{ \frac{|Y_i|}{S} \leq c_i \right\}. \tag{1}
\]

In words, the probability that a random normal vector (divided by \( S \)) with arbitrary covariances falls inside a rectangle centered at its mean is always at least as large as the corresponding probability for the case where the covariances are zero, i.e., where the coordinate variables are independent.

Application of this inequality to confidence intervals or tests for multiple \( t \) statistics which may have dependent numerators is immediate. It produces slightly sharper intervals or tests than the Bonferroni inequality because \( (1-\alpha)^r > 1 - r\alpha \). Whether for use with the Šidák inequality or the original Bonferroni inequality, recent charts by Moses [81] facilitate determination of the appropriate \( t \) critical value.

Earlier Dunn [1] had proved this inequality for \( k = 2 \) or 3 or for general \( k \) but special \( \Sigma \). Also, in earlier work Slepian [6] had obtained a similar one-sided inequality. Slepian's inequality states
that if $\sigma_{ij}^2 \geq \sigma_{ij}^2$ for $i \neq j$ and $\sigma_{ii}^2 = \sigma_{ii}^2$ for $i = 1, \ldots, r$, then

$$P_1\{Y_1 \leq c_1, \ldots, Y_r \leq c_r\} \geq P_2\{Y_1 \leq c_1, \ldots, Y_r \leq c_r\}, \quad (2)$$

where $P_k$, $k = 1, 2$, is computed under the multivariate distribution $N(0, \Sigma_k)$ with $\Sigma_k = (\sigma_{ij}^k)$.

For a summary of results on the case where the denominator variables $S$ in (1) is allowed to be different for different $Y_i$ the reader is referred to Sidák [26], Khatri [19], [20] considered generalizations of $P\{\bigcap_1^r \{Y_i \leq c_i\} \geq \bigcap_1^r \{Y_i \leq c_i\} \geq \bigcap_1^r \{Y_i \leq c_i\}$ to symmetric convex regions about the origin as well as the corresponding inequality $P\{\bigcap_1^r \{Y_i \geq c_i\} \geq \bigcap_1^r \{Y_i \geq c_i\} \geq \bigcap_1^r \{Y_i \geq c_i\}$ for the complementary sets.

2.2 Methods for Unbalanced ANOVA

At the time of writing of [6] the Tukey studentized range technique was applicable only for the case of equal sample sizes.

For one-way classifications where different populations had different samples sizes the only techniques available were Bonferroni t statistics or Scheffé F projection intervals. In practice the fearless used the studentized range with a modal or median sample size for slightly imbalanced designs.

Recently, a couple new methods have been proposed. For $Y_{ij}$ distributed independently as $N(\mu_i, \sigma^2)$, $i = 1, \ldots, r$, $j = 1, \ldots, n_i$,

Spjotvoll and Stoline [53] gave the intervals

$$\mu_i - \mu_j \leq \bar{Y}_i - \bar{Y}_j \pm q_{\alpha, r, \nu} \max \left( \frac{1}{\sqrt{n_i}}, \frac{1}{\sqrt{n_j}} \right), \quad (3)$$

5
where $\bar{Y}_i = \frac{\sum_i Y_{ij}}{n_i}$, $S^2 = \frac{\sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2}{\nu}$, $\nu = \frac{r}{r-1}(n_i-1)$, and $q'_{\alpha, r, \nu}$ is the $1 - \alpha$ percentile point of the studentized augmented range distribution (cf., [6, p. 40]). These intervals are conservative in the sense that the probability they simultaneously contain the true mean differences for all $i, j$ is greater than or equal to $1 - \alpha$. Tables of the studentized augmented range distribution are not available, but the critical point $q_{\alpha, r, \nu}$ from a studentized range distribution (cf., [6, p. 38]) gives a good approximation to $q'_{\alpha, r, \nu}$ provided $r > 2$ and $\alpha < .05$.

Hochberg [49] gave the different intervals

$$
\mu_i - \mu_j \in [\bar{Y}_i - \frac{1}{2} |m|_{\alpha, c, \nu} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}],
$$

(4)

where $c = \left\lceil \frac{r}{2} \right\rceil$ and $|m|_{\alpha, c, \nu}$ is the $1 - \alpha$ percentile point of the studentized maximum modulus distribution (cf., [6, p. 71]). Use of this technique is facilitated by improved tables of critical points for the studentized maximum modulus distribution computed by Hahn and Hendrickson [34].

There arises the natural question of which technique gives the shortest intervals. Ury [62] has made a valuable study of the Spjøtvoll-Stoline and Hochberg procedures as well as the Scheffé S-method and the Bonferroni intervals as modified by Šidák [24] and Dunn [48]. Unfortunately, the answer is not simple. No one of the four techniques dominates the others. Which type of interval is best depends upon the particular combination of sample sizes, significance level, number of populations, and degrees of freedom. Roughly, the
Spjøtvoll-Stoline method (3) is preferable for very mildly imbalanced designs. For designs with greater imbalance the Hochberg procedure (4) is best when the number of populations is 6 or smaller and the Bonferroni-Dunn-Sidak intervals are best for 7 or more populations. Ury's paper contains two tables which are very helpful in determining the region of optimality for each technique.

Two noteworthy papers have appeared which compare some of the aforementioned procedures with other multiple comparisons methods for the one-way classification such as multiple range tests, least significant difference tests, and empirical Bayes tests. The papers are by Carmer and Swanson [56] and by Einot and Gabriel [57]. A conclusion of the first paper is that for the testing problem protected least significant difference tests (cf., [6, pp. 90-94]) and empirical Bayes tests (cf., Section 2.4) are superior in terms of power with the former being simpler to use and explain in practice. The second paper favors the use of Tukey's studentized range technique for simplicity and little loss in power.

For recent work on the companion problem of unequal population variances $\sigma_i^2$, i = 1, ..., r, the reader is referred to Hochberg [51] and Tamhane [54].

2.3. **Conditional Confidence Levels**

Olshen [77] proved the following interesting result under mild conditions on n, p, and $\alpha$:

$$ P \left\{ \frac{1}{p} \frac{(\hat{\beta} - \beta)^T (X^T X) (\hat{\beta} - \beta)}{\sigma^2} < F_{\alpha, p, n-p} \left| \frac{1}{p} \frac{(\hat{\beta}^T (X^T X) \hat{\beta}}{\sigma^2} > F_{\alpha, p, n-p} \right\} \right\} < 1 - \alpha \quad (5) $$
for all $\beta$ and $\sigma^2$ where $Y \sim N(X\beta, \sigma^2 I)$, $X$ is an $n \times p$ matrix of full
rank, $\hat{\beta} = (X^T X)^{-1} X^T Y$, $\hat{\sigma}^2 = \frac{Y^T (I - X (X^T X)^{-1} X^T) Y}{(n-p)}$, and $F_{\alpha, p, n-p}$
is the $1 - \alpha$ percentile point of an $F$ distribution with $p$ and $n-p$
degrees of freedom. This means that if Scheffe simultaneous confidence intervals are computed only when the $F$ test is significant,
then the conditional probability of coverage for the confidence intervals is always less than the nominally stated unconditional
probability. Monte Carlo experiments indicate the discrepancy can be substantial.

Similar effects on conditional error rates for other multiple comparisons procedures have been established by Bernhardson [76].

2.4. Bayes Approach

In a series of papers ([85], [86], [87], [88], [89]) Duncan
and his associates have developed the Bayesian approach to multiple comparisons in a balanced one-way classification.

Specifically, they assume an additive linear loss model. For a component problem involving a decision of whether $\mu_i > \mu_j$ or $\mu_i \leq \mu_j$ the linear loss structure is

$$L(d^+, \Delta) = \begin{cases} k_1 |\Delta|, \Delta \leq 0 \\ 0, \Delta > 0 \end{cases}$$

and

$$L(d^0, \Delta) = \begin{cases} 0, \Delta \leq 0 \\ k_2 \Delta, \Delta > 0 \end{cases}$$

where $\Delta = \mu_i - \mu_j$, $d^+$ is the decision $\Delta > 0$, and $d^0$ is the decision $\Delta \leq 0$. The ratio $k = k_1/k_2$ measures the relative seriousness of
type I errors with respect to type II errors. The overall loss is
the sum of the component losses over the decisions \( \{ \mu_i > \mu_j \} \) vs. \( \{ \mu_i \leq \mu_j \} \) and \( \{ \mu_j > \mu_i \} \) vs. \( \{ \mu_j \leq \mu_i \} \) for all pairs \( i, j \).

The usual normal theory model \( Y_{ij} \sim N(\mu_i, \sigma_e^2), i = 1, \ldots, r, \]
\( j=1, \ldots, n, \) is adopted for the observations. The prior distribution on the \( \mu_i \) assumes they are independently distributed as \( N(0, \sigma^2_{\mu}) \), and a truncated product of two independent conjugate \( \chi^2 \) densities is assumed for \( \sigma_e^2 \) and \( \sigma_a^2 = \sigma_e^2 + n \sigma_{\mu}^2 \).

Because of the additivity of the loss structure the overall risk is minimized by minimizing each component risk, and this leads to the \( k \)-ratio \( t \) test which is to decide

\[
\mu_i > \mu_j \quad \text{if} \quad \overline{Y}_{i*} - \overline{Y}_{j*} > t(k, F, q, f) S_d,
\]
\[
\mu_i = \mu_j \quad \text{if} \quad \mid \overline{Y}_{i*} - \overline{Y}_{j*} \mid < t(k, F, q, f) S_d,
\]
\[
\mu_i < \mu_j \quad \text{if} \quad \overline{Y}_{i*} - \overline{Y}_{j*} < -t(k, F, q, f) S_d.
\]

The critical \( t \) value is a function of the loss ratio \( k = k_1/k_2 \), the Bayesian F ratio

\[
F = \frac{f}{q} \cdot \frac{n \sum_{i=1}^{r} (\overline{Y}_{i*} - \overline{Y}_{.})^2 + q_p \sigma_a^2}{\sum_{i=1}^{r} \sum_{j=1}^{n} (Y_{ij} - \overline{Y}_{i*})^2 + f_p \sigma_e^2}
\]

where \( \sigma_e^2 \) and \( \sigma_a^2 \) are the prior values for \( \sigma_e^2 \) and \( \sigma_a^2 = \sigma_e^2 + n \sigma_{\mu}^2 \), and the Bayesian degrees of freedom

\[
f = r(n-1) + f_p,
\]
\[
q = (r-1) + q_p
\]
where \( f_p \) and \( q_p \) are the degrees of freedom for the conjugate \( \chi^2 \) prior distributions. Tables of \( t(k,F,q,f) \) are available in the Corrigenda of [87]. The Bayesian standard deviation estimate \( S_d \) used in (7) is the square root of

\[
S_d^2 = \frac{2}{n} \cdot \frac{\sum_i^n \sum_j^1 (\bar{Y}_{ij} - \bar{Y}_{i*})^2}{f} + f_p \frac{\bar{c}_e^2}{\sigma_e^2}.
\]

(10)

The innovative feature of this approach is the adaptive nature of the decision rule. How large each difference \( |\bar{Y}_{i*} - \bar{Y}_{j*}| \) must be in order for the population means to be declared significantly different depends upon the F ratio of between to within population variation. The critical value \( t(k,F,q,f) \) is a decreasing function of \( F \) so the larger the F ratio for the test of the equality of all the means, the easier it is to reject the equality of an individual pair of means. Thus, this approach accomplishes in a smooth fashion what the Fisher protected LSD test attempts to do in a two stage manner.

Confidence intervals associated with the above k-ratio \( t \) test have been derived in [85]. In the large sample case \( (q=f=\infty) \) they take the form

\[
\mu_i - \mu_j \in (1 - \frac{1}{F})(\bar{Y}_{i*} - \bar{Y}_{j*}) \pm (1 - \frac{1}{F})^{1/2} S_d t(k),
\]

(11)

where \( t(k) \) is tabled in [85]. Dixon and Duncan have established that these intervals are large sample approximations to the minimum Bayes risk intervals for additive squared-error loss.

The intervals (11) provide a link between the theories of simultaneous confidence intervals and simultaneous estimation. The
form of the point estimator in (11) is nearly identical to the James-Stein estimator for a normal mean in a one-way classification

\[ \hat{\mu}_i = \overline{Y} + \left( 1 - \frac{r - 3}{r(n-1) + 2} \cdot \frac{\sum_{i=1}^{r} \sum_{j=1}^{n} \left( y_{ij} - \overline{Y} \right)^2}{n \sum_{i=1}^{r} \left( \overline{Y} - \overline{Y} \right)^2} \right) (\overline{Y} - \overline{Y}) \]. (12)

Efron and Morris [2] have given an empirical Bayes interpretation to the James-Stein estimator under additive squared-error loss which ties in with the approach of Dixon and Duncan.

For additional work on Bayes confidence sets with implications for multiple comparisons problems the reader is referred to Faith [3].

2.5. Confidence Bands in Regression

For the simple linear regression model \( Y_i = \alpha + \beta(x_i - \overline{x}) + e_i \) with \( e_i \sim N(0, \sigma^2) \) the principal confidence band available in 1966 for bounding the regression line \( \alpha + \beta(x - \overline{x}) \) was the Working-Hotelling-Scheffé hyperbolic band (cf., [6, pp. 110-114]). Since then, a plethora of confidence bands have appeared with many different shapes and many different purposes.

Bowden and Graybill [132] extended and simplified the earlier work of Gafarian [4] on confidence bands of uniform width over an interval \( x \in [a, b] \). The uniform width confidence band is

\[ \alpha + \beta(x - \overline{x}) \in \hat{\alpha} + \hat{\beta}(x - \overline{x}) \pm c^*_\alpha \hat{\sigma}, \ x \in [a, b], \] (13)

where \( \hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 \) are the usual least squares estimators. The critical constant \( c^*_\alpha \) is computable from the bivariate t distribution, but
depends on \( a, b, n, \bar{x}, \sum_{i=1}^{n}(x_i-\bar{x})^2 \) as well as the significance level \( \alpha \). Bowden and Graybill give tables of \( c^2_\alpha \) which, with interpolation, will cover a wide range of cases. For applications outside the scope of the tables conservative bands can be achieved by constructing a confidence rectangle for

\[
\alpha + \beta(a-\bar{x}) \quad \text{and} \quad \alpha + \beta(b-\bar{x})
\]

(14)

from the Bonferroni or Šidák inequalities applied to the univariate \( t \) distribution. Trapezoidal bands could be obtained instead by allowing the confidence intervals placed on the two quantities in (14) to vary in width.

Graybill and Bowden [135] derived a linear segment confidence band

\[
\alpha + \beta(x-\bar{x}) \in \hat{\alpha} + \hat{\beta}(x-\bar{x}) \pm |m|_{\alpha,2,n-2} \hat{\sigma} \left[ \frac{1}{\sqrt{n}} + \frac{|x-\bar{x}|}{\left( \sum_{i=1}^{n}(x_i-\bar{x})^2 \right)^{1/2}} \right]
\]

(15)

for \(-\infty < x < +\infty\). The band (15) is obtained from projections of a rectangular confidence region for \( \alpha \) and \( \beta \) based on the independent normal estimates \( \hat{\alpha} \) and \( \hat{\beta} \). This is why the \( 1 - \alpha \) percentile point \( |m|_{\alpha,2,n-2} \) of a studentized maximum modulus distribution on \( 2 \) and \( n-2 \) degrees of freedom enters as the critical constant. Again, the improved tables for this distribution by Hahn and Hendrickson [34] are useful. The Graybill-Bowden paper improves on the work of Folks and Antle [134] who derived conservative linear segment bands.

Halperin and Guarian [136] studied the distribution theory when the Working-Hotelling-Scheffé hyperbolic band is constrained to
a finite interval \( x \in [a,b] \). For an interval symmetric about \( \bar{x} \) the special tables which are required for the determination of the appropriate critical constant appear in an earlier work [137]. Although the Halperin-Guarian band has smaller average width than the Bowden-Graybill uniform width band, the latter band and the Graybill-Bowden linear segment band have the advantage of tables which cover a broader range and are easier to use. Both linear bands also enjoy the convenience of being easier to graph.

Dunn [133] modified the Graybill-Bowden linear segment band by considering its restriction to a finite interval \( x \in [a,b] \) symmetric about \( \bar{x} \). A factor involving the length of the interval \([a,b]\) enters into the expression (15) and the critical constant comes from a studentized maximum modulus distribution with 3 degrees of freedom in the numerator.

For the quadratic regression model \( \beta_0 + \beta_1 x + \beta_2 x^2 \) the only tool for constructing a confidence band until recently has been the F projection method of Scheffé (cf., [6, pp. 110-114]). This band is conservative because the vectors \((1,x,x^2)\) do not map out a full three-dimensional space as \( x \) varies from \(-\infty\) to \(+\infty\). Wynn and Bloomfield [141] studied how to sharpen the Scheffé-type band and provided some tables for doing so. Trout and Chow [139], on the other hand, considered a uniform width band restricted to \( x \in [a,b] \). Unfortunately, tables for applying this procedure are rather limited.

Bohrer [125] studied sharpening of the Scheffé band in the multiple regression model \( \hat{x} = (x_1, \ldots, x_p)(\hat{\beta}_1, \ldots, \hat{\beta}_p)^T \) where \( \hat{x} \) is restricted to the non-negative orthant \( \mathbb{R}_+^p = \{x|x_i \geq 0, i = 1, \ldots, p\} \).
There may be instances in which only a one-sided confidence band is required for the regression surface. Presumably, then, one could improve over the two-sided band. Bohrer [126] proved that this cannot be accomplished for a Scheffé-type band in a multiple regression model $x_\beta$ with no intercept (i.e., no $x_i = 1$). However, for $x$ restricted to a subset of the non-negative orthant $R_+$ improvement is possible and is sometimes substantial as demonstrated by Bohrer and Francis [129]. Also, Bohrer and Francis [130] showed how to obtain a sharper one-sided band for simple linear regression $\alpha + \beta(x-x)$ when $x$ is confined to an interval $[a,b]$. Finally, Hochberg and Quade [138] gave a sharpened one-sided band for multiple regression $\beta_0 + x_\beta$ with an intercept $\beta_0$.

3. References


4.1. Survey Articles


4.2. Probability Inequalities


4.3. Tables


4.4. Normal Multi-factor Methods

Unbalanced and Heteroscedastic ANOVA


**Comparison of Procedures**


Robustness


**Power Functions**


**Conditional Error Rates**


**Graphical Techniques**


Bayes Techniques


Several Treatments vs. Control


Tests on Interactions


Tests on Variances


General


4.5. **Regression**

**Confidence Bands**


Calibration and Prediction


General


4.6. Categorical Data


4.7. Nonparametric Techniques


### 4.8 Multivariate Methods


4.9. **Miscellaneous**


designs and large-sample tests for linear hypotheses.
Biometrika 62, 71-78.

intervals for parametric functions from a parametric con-

tests for trend and serial correlations for Gaussian
Markov residuals. Econometrica 34, 472-480. Erratum: 34
(1966), 908-909.

[214] Levy, K. J. (1975). Large-sample pair-wise comparisons
involving correlations, proportions, or variances.
Psychological Bulletin 82, 174-176.

[215] Levy, K. J. (1975). Large-sample many-one comparisons
involving correlations, proportions, or variances.
Psychological Bulletin 82, 177-179.

[216] Likeš, J. (1968). Note on Tukey's method of multiple

in samples from an exponential distribution. Biometrische
Zeitschrift 15, 545-555.

[218] Mantel, N. (1968). Simultaneous confidence intervals and
experimental design with normal correlation. Biometrics
24, 434-437.

Psychological Bulletin 65, 280-290.

parameters. IEEE Transactions on Reliability R-24,
186-192.

classifying all pairs out of k means as close or distant.
Journal of the American Statistical Association 70,
832-838.

[222] Rodger, R. S. (1967). Type I errors and their decision basis.
British Journal of Mathematical and Statistical
Psychology 20, 51-62.


4.10.  Late Additions
