JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 73B, No. 4, October–December 1969

The Partitioning of Interaction in Analysis of Variance

John Mandel
Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234

(June 2, 1969)

A method is presented for the analysis of data representing functions of two variables, when the response can be tabulated in a rectangular array. The procedure is based on a partitioning of the row by column interaction effects into a sum of terms, each of which is the product of a row factor by a column factor. The factors in each term are estimated by a method involving the extraction of characteristic roots.

The method contains as special cases a number of procedures used for the handling of non-additivity in two way arrays. It is very useful for the fitting of empirical surfaces, but is also applicable to cases in which the data depend on qualitative rather than quantitative factors.

Comparisons with other techniques are made and an illustrative example is given.

Key words: Factorial Experiments; interaction; nonadditivity; principal components; surface fitting.

1. Introduction

In a previous paper [8],[1] a method was presented for the analysis of data representing functions of two variables, when the response can be tabulated in a rectangular array. The analysis was based on the assumption of a "linear model," the validity of which was subject to verification in each individual case. Essentially, the linear model assumption is that the elements of each row, when plotted against the column averages of the table (or that the elements of each column, when plotted against the row averages), provide a straight line. It was pointed out in [8] that when the linear model assumption does not hold, the analysis can be handled in two ways. The first of these simply generalizes the assumption of "linearity," required by the linear model, to one of quadratic behavior, as discussed in reference [9].

The second way of dealing with more complex models than are covered by the linearity assumption is to attack the problem of row by column interaction in a completely systematic way, by partitioning this term in as many individual terms as are required by the data. Thus, no prior assumptions such as linearity, or concurrence [7], or quadratic behavior [9] are made; to a considerable extent the data themselves generate the model. It is this approach that we wish to discuss in the present paper.

We will assume that the response is a quantitative variable, generally a measure of some property of a material or of a system. The nature of the data, and of the problem behind the data then depends on the nature of the two independent variables (which are represented by the rows and the columns of the two-way table). These may be qualitative or quantitative, or mixtures of both types, and the interpretation of the analysis will depend on which of these situations pertains. For the sake of brevity, we will describe the application of the method to only one illustrative example. Other interesting applications of the method will be presented in subsequent papers.

[1] Figures in brackets indicate the literature references at the end of this paper.

A basic feature of the method discussed in the previous publication [8] was the division of the analysis in two phases, the first of which was referred to as the analysis for "internal structure." This first phase led to an expression relating the response to sets of parameters, each of which was a function of only one variable. The second phase of the analysis consisted in relating these parameters to the original row and column variables. The present analysis, while far more general, retains this basic feature.

In the course of our discussion we will call attention to the relation between the proposed approach and other models for two-way data. We will also discuss the similarities, as well as the differences, between this approach and some related procedures discussed in the literature. In a broad sense, all these procedures, including the one presented in this paper, derive from the "method of principal components," the development of which is due primarily to Harold Hotelling [4]. Thus, while the method presented in this paper rests largely on well-known mathematical results, the approach is nevertheless novel. In particular, while practically all discussions of this method of principal components are nonsymmetrical with respect to the rows and the columns of the two-way table, the present approach treats the rows and the columns in the same identical way.

Ideas similar to those presented here have been discussed by E. J. Williams [17] and by Pike and Silverberg [11].[2] Williams makes explicit use of the theory of latent roots underlying the method of principal components, but approaches the data analysis problem from a somewhat different point of view. Pike and Silverberg make no use of the method of principal components.

While this paper was being prepared for publication, a most interesting paper by Gollob [3] was published. It is apparent that we developed the same basic model independently of each other.

A basic difference between Gollob's method and mine lies in what I believe to be a novel approach to the question of degrees of freedom in principal component analysis. This approach is discussed in detail in the present paper.

2. The Analysis of Internal Structure

Let x_i and y_j denote the levels of the two independent variables or categories, represented by the rows and columns of the table and z_{ij} the observed value of the measured (dependent) variable for x_i and y_j.

Assume that the table consists of m rows and n columns. We begin by writing the model usually adopted in analysis of variance:

$$z_{ij}=\mu+\rho_i+\gamma_j+\eta_{ij} \qquad \begin{cases} i=1 \text{ to } m \\ j=1 \text{ to } n \end{cases}. \tag{1}$$

Here, μ is a constant, estimated by the grand-mean; ρ_i is the "row main effect" and γ_j the "column main effect"; the quantity η_{ij} represents the "interaction between row i and column j."

As usual we impose the conditions:

$$\sum_i \rho_i=0 \qquad \sum_j \gamma_j=0 \qquad \sum_i \eta_{ij}=\sum_j \eta_{ij}=0. \tag{2}$$

It should be noted that eq (1) expresses the function of two variables, z_{ij}, *partly* in terms of two functions of a single variable each, ρ_i and γ_j, but that it also involves a new function of two variables, η_{ij}.

It is often assumed that η_{ij} is a random variable, of mean zero, and standard deviation σ. In that case, which is known as the "additive case," eq (1) constitutes a real simplification: the replacement of a function of two variables by two functions of one variable each.

[2] For calling my attention to the latter reference I am indebted to Professor David L. Wallace.

When the assumption of additivity is not definitely known to be valid, the analysis can proceed only on the basis of some other definite assumption concerning the structure of η_{ij}. The assumption we will adopt in this paper is expressed by the following equation:

$$\eta_{ij}=\theta u_i v_j+\theta' u_i' v_j'+\theta'' \mu_i'' v_j'' \ldots +\epsilon_{ij} \tag{3}$$

where ϵ_{ij} is a random variable of zero mean and standard deviation σ. When η_{ij} is expressed by only a few terms of the multiplicative type $\theta u_i v_j$, eq (3) also constitutes a real simplification. Without the loss of generality, we may impose the constraints:

$$\sum_i u_i=\sum_i u_i'=\ldots=\sum_j v_j=\sum_j v_j'=\ldots=0 \tag{4}$$

$$\sum_i u_i^2=\sum_i u_i'^2=\ldots=\sum_j v_j^2=\sum_j v_j'^2=\ldots=1. \tag{5}$$

The problem is to find estimates for the new parameters $\theta, \theta', \ldots u, u', \ldots, v, v', \ldots$ and for the standard deviation of random error, σ. Equations (1) and (3) together with the constraints expressed by eqs (4) and (5), constitute our model. We will refer to the quantities μ, ρ_i, γ_j, θ, u_i, v_j, θ', u_i', v_j', etc. as "structural parameters."

3. Solution by Least Squares [3]

The interaction η_{ij} is estimated by the residual

$$d_{ij}=z_{ij}-\hat{\mu}-\hat{\rho}_i-\hat{\gamma}_j \tag{6}$$

where $\hat{\mu}$, $\hat{\rho}_i$, $\hat{\gamma}_j$ are the usual estimates

$$\hat{\mu}=z_{\cdot\cdot};\ \hat{\rho}_i=z_{i\cdot}-z_{\cdot\cdot};\ \hat{\gamma}_j=z_{\cdot j}-z_{\cdot\cdot} \tag{7}$$

The dot notation indicates here, as usual, averaging over the subscript replaced by the dot. To fit the interaction, consider first the model:

$$\eta_{ij}=\theta u_i v_j+\epsilon_{ij} \tag{8}$$

subject to the constraints expressed by eqs (4) and (5). We obtain estimates for θ, u_i, and v_j by minimizing with respect to θ, u_i, and v_j, the quantity:

$$S=\sum_i \sum_j (d_{ij}-\theta u_i v_j)^2-\lambda_1(\Sigma u^2-1)-\lambda_2(\Sigma v^2-1)-2\mu_1\Sigma u_i-2\mu_2\Sigma v_j \tag{9}$$

where the λ and μ's are Lagrange multipliers. Carrying out the calculations, one obtains the relations:

$$u_k=\theta^{-1}\sum_j d_{kj}v_j \tag{10}$$

$$v_l=\theta^{-1}\sum_i d_{il}u_i \tag{11}$$

[3] The derivation in this section is essentially the same as that given by Williams [17].

Eliminating v_j between (10) and (11), and setting

$$s_{ik} = \sum_j d_{ij} d_{kj} \tag{12}$$

one obtains:

$$\theta^2 u_k = \sum_i u_i s_{ik}. \tag{13}$$

In matrix notation, eqs (12) and (13) become:

$$(s) = (d)(d)^T \tag{14}$$

and

$$\theta^2 (u) = (s)(u) \tag{15}$$

where T indicates a transpose, (s) is an $m \times m$ matrix, and (u) is an $m \times 1$ vector.

It follows from (15) that θ^2 is an eigenvalue of the matrix (s) and that (u) is an associated eigenvector.

Similarly it can be shown that:

$$\theta^2 (v) = [(d)^T (d)](v). \tag{16}$$

where (v) is an $n \times 1$ vector.

The desired minimization of S (eq (9)) is accomplished by taking the *largest* eigenvalue θ^2 of the matrix (s). Furthermore, an important consequence, which can readily be proved,[4] is that if θ^2 is chosen to be the *largest* eigenvalue of the matrix (s), and one attempts to find the values of θ', u_i' and v_j' such that:

$$\sum \sum_j [(d_{ij} - \theta u_i v_j) - \theta' u_i' v_j']^2 \tag{17}$$

is a minimum, subject to the constraints (4) and (5), it turns out that θ'^2 is simply the second-largest eigenvalue of the same matrix (s), and (u') the associated eigenvector, with a similar situation for (v').

Thus, by obtaining the *complete* set of eigenvalues of (s), and the associated sets of vectors (u), (v), (u'), (v'), etc., one actually obtains the least-squares solution of *all* parameters in eq (3).

4. An Analog to Analysis of Variance

In terms of sample estimates, eq (3) leads to the equation:

$$d_{ij} = \hat{\theta} \hat{u}_i \hat{v}_j + \hat{\theta}' \hat{u}_j' \hat{v}_j' + \ldots + e_{ij}. \tag{18}$$

The rank of the (s) matrix satisfies the inequality:

$$\text{rank (s)} \leq \min (m, n) - 1. \tag{19}$$

Consequently, the number of terms on the right side of (18) is at most min $(m, n) - 1$. In analogy to the usual interpretation of the results of analysis of variance, a decision will first be made (see following section) on how many terms of the type $\theta u_i v_j$ should be retained in the model; the residual sum of squares is then used for an estimation of the variance of ϵ.

[4] See, for example: Harman, H. H., Modern Factor Analysis (The University of Chicago Press, Chicago, 1960, chap. 9).

Assume that we calculate *all* terms of the type $\theta u_i v_j$ so that ϵ_{ij} is taken equal to zero. Then we obtain at once:

$$\sum_i \sum_j d_{ij}^2 = \hat{\theta}^2 + \hat{\theta}'^2 + \hat{\theta}''^2 + \ldots . \tag{20}$$

This follows from eq (18) and the orthogonality property of eigenvectors, by which all cross-products are zero.

Equation (20) constitutes an actual partitioning of the sum of squares of interaction. One is therefore tempted to try an approach of the analysis of variance type. However, the θ^2 are *not* quadratic forms in the original measurements z_{ij}. Nevertheless, it is possible to formulate the problem in analysis of variance language, by virtue of the following considerations.

Let $\mathcal{N}(\mu, \sigma)$ represent a normal population of mean μ and standard deviation σ.

Suppose first that the z_{ij} are a random sample from a normal population $\mathcal{N}(0, 1)$. Then, the quantities $\hat{\theta}^2$, $\hat{\theta}'^2$, $\hat{\theta}''^2$, etc., each form a definite statistical population. Let:

$$M_1 = E(\hat{\theta}^2)_{\mathcal{N}}, \; M_2 = E(\hat{\theta}'^2)_{\mathcal{N}}, \; M_3 = E(\hat{\theta}''^2)_{\mathcal{N}}, \ldots \tag{21}$$

where E represents an expected value and the symbol $\mathcal{N}$ refers to the standard normal distribution $\mathcal{N}(0, 1)$.

If the normal population in question had a variance of σ^2, instead of unity, the quantities M_1, M_2, M_3, etc., would simply be multiplied by σ^2.

Therefore, the ratios of the $\hat{\theta}^2$ values obtained from a random sample from $\mathcal{N}(0, \sigma)$ to the corresponding M values are *all* estimates of σ^2.

If, now, a matrix of data z_{ij} is given, and if the interaction terms are nothing but random gaussian error, then the ratios of the $\hat{\theta}^2$ obtained from these data to the corresponding M-values obtained from an $\mathcal{N}(0, 1)$ matrix of the same dimensions are simply estimates of σ^2.

Thus, the M-values fulfill a role that is quite analogous to that of the degrees of freedom in ordinary analysis of variance. More specifically, the M are such that, for random gaussian interaction, the ratio of each $\hat{\theta}^2$ by the corresponding M is an unbiased estimate of σ^2.

Carrying the above argument a little further, and using heuristic reasoning, we may expect that if the real model contains say, k terms of the type $\theta u_i v_j$, then the corresponding k values of $\hat{\theta}^2$ will be inflated by the systematic effects of these terms, while the *remaining* terms, say:

$$\hat{\theta}_{k+1}^2, \hat{\theta}_{k+2}^2, \text{ etc.}$$

will only be estimates of

$$M_{k+1}\sigma^2, \; M_{k+2}\sigma^2, \text{ etc.}$$

Thus:

$$\hat{\theta}_{k+1}^2/M_{k+1}, \; \hat{\theta}_{k+2}^2/M_{k+2}, \text{ etc.},$$

will all be estimates of σ^2.

In this way a judgment can be made as to the number of terms, k, that should be retained in the model, as will be shown in the discussion of the example, further in this paper.

5. A Monte-Carlo Study

To verify the above assumptions, and obtain reliable numerical estimates for the M-values, a series of Monte Carlo experiments were carried out, as follows.

For specific values of m and n, matrices containing random normal deviates from $\mathcal{N}(0, 1)$, were analyzed by the method outlined above. For each m and n combination, 625 such matrices were analyzed, and the estimates of the corresponding M-values computed. The averages of the θ^2, for 625 sets, and their standard deviations are listed in tables A1 and A2, respectively, of the appendix.

362-120 O - 69 - 3

In an $m \times n$ matrix, the number of degrees of freedom for interaction is $(m-1)\ (n-1)$. If $\sigma^2=1$, the expected value of the sum of squares for interaction is also $(m-1)\ (n-1)$. Thus, the expected value of the sum of the M-values, for any given matrix of size $m \times n$, is $(m-1)\ (n-1)$. In view of this fact, it seemed advantageous to compute the ratio of each value to $(m-1)\ (n-1)$. These ratios, expressed as percent, are denoted "Percent of Total Interaction Degrees of Freedom," and are listed in table A3. They were obtained by a weighted least squares adjustment process, using the reciprocals of the squares of the observed standard deviations as relative weights, and introducing the constraint that the sum of the percentages be 100. The values in table A3 lend themselves to more precise interpolation than those of table A1.

Additional Monte Carlo studies were carried out to verify the conjecture that if one or more terms of the type $\theta u_i v_j$ were actually present in the true model, the *subsequent* terms would not be affected by these systematic terms. The experiments showed that the introduction of even very large systematic terms of the type $\theta u_i v_j$ had the effect of only slightly inflating the subsequent terms; these remained, for all practical purposes, acceptable estimates of the experimental error.

6. Practical Use of the Method

Let z_{ij} be given in the form of an $m \times n$ matrix. We assume that z_{ij} is a function of x and y, where the values of x correspond to the rows and the values of y to the columns of the matrix.

We first fit a model of the type:

$$z_{ij}=\mu+\rho_i+\gamma_j+\eta_{ij}=\mu+\rho_i+\gamma_j+\theta u_i v_j+\theta' u_j' v_j'+ \ . \ . \ .$$

to the data, ignoring momentarily the numerical values, if any, of x_i and y_j.

The usual analysis of variance yields the estimates:

$$\hat{\mu}=z_{..};\ \hat{\rho}_i=z_{i.}-z_{..};\qquad \hat{\gamma}_j=z_{.j}-z_{..}$$

and the residuals:

$$\hat{\eta}_{ij} \equiv d_{ij}=z_{ij}-\hat{\mu}-\hat{\rho}_i-\hat{\gamma}_j=z_{ij}+z_{..}-z_{i.}-z_{.j}.$$

For purposes which will become apparent in the discussion of the illustrative example, we "standardize" the parameter estimates $\hat{\rho}_i$ and $\hat{\gamma}_j$, by dividing each of them by the square root of the corresponding sum of squares. Thus, writing:

$$R=\sqrt{\sum_i \hat{\rho}_i^2},\qquad G=\sqrt{\sum_j \hat{\gamma}_j^2}$$

$$r_i=\frac{1}{R}\hat{\rho}_i,\qquad c_j=\frac{1}{G}\hat{\gamma}_j \tag{22}$$

our model becomes:

$$z_{ij}=\mu+Rr_i+Gc_j+\theta u_i v_j+\theta' u_i' v_j'+ \ . \ . \ . \ . \tag{23}$$

Applying a diagonalization technique to the matrix $(d) \cdot (d)^T$, we find:

(a) a set of eigenvalues

$$\hat{\theta}^2,\ \hat{\theta}'^2,\ \hat{\theta}''^2,\ . \ . \ .$$

(b) a set of u vectors, each of dimensions $m \times 1$:

$$(u),\ (u'),\ (u''),\ . \ . \ .$$

(c) a set of v vectors, each of dimensions $n \times 1$:

$$(v),\ (v'),\ (v''),\ . \ . \ .$$

We now tabulate the results as follows (tables 1 and 2):

TABLE 1. *Row-dependent parameters*

Row variable	Row effects, r_i	u_i	u_i'	u_i'' . . .
x_1	r_1	u_1	u_1'	u_1'' . . .
x_2	r_2	u_2	u_2'	u_2'' . . .
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_m	r_m	u_m	u_m'	u_m'' . . .

TABLE 2. *Column-dependent parameters*

Column varaible	Column effects, c_j	v_j	v_j'	v_j'' . . .
y_1	c_1	v_1	v_1'	v_1'' . . .
y_2	c_2	v_2	v_2'	v_2'' . . .
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
y_n	c_n	v_n	v_n'	v_n'' . . .

We form the analysis of variance table (table 3):

TABLE 3. *Analysis of variance* [a]

Source	*DF*	*SS*	$MS=\frac{SS}{DF}$
Total	$m \cdot n$	$\sum_i \sum_j z_{ij}^2$	
μ	1	$mn\mu^2$	
ρ_i	$m-1$	nR^2	$nR^2/(m-1)$
γ_j	$n-1$	mG^2	$mG^2/(n-1)$
η_{ij}	$(m-1)(n-1)$	$\sum_i \sum_j d_{ij}^2$	$\sum \sum d_{ij}^2/(m-1)(n-1)$
$\theta u_i v_j$	M_1	$\hat{\theta}^2$	$\hat{\theta}^2/M_1$
$\theta' u_i' v_j'$	M_2	$\hat{\theta}'^2$	$\hat{\theta}'^2/M_2$
$\theta'' u_i'' v_j''$	M_3	$\hat{\theta}''^2$	$\hat{\theta}''^2/M_3$
.	.	.	.
.	.	.	.
.	.	.	.

[a] The term "analysis of variance" is used here in a generalized sense: as in classical analysis of variance, the sum of squares is partitioned and mean squares are calculated. Under the null-hypothesis of "no effect," the mean squares are unbiased estimates of the error variance. The *M*-values are "degrees of freedom" only in the sense of appropriate divisors for the corresponding sums of squares, as explained in the text.

We now examine the mean squares corresponding to the breakdown of the η_{ij} into the sum of multiplicative terms $\theta u_i v_j$, $\theta' u_i' v_j'$, etc. At the present time, no distribution theory is available for these mean squares, and exact tests of significance cannot be carried out. Nevertheless, even an intuitive appraisal of the mean squares generally leads to fairly clear-cut decisions concerning the number of product terms that should be retained in the model. This will be illustrated in the next section, which deals with an illustrative example.

7. Illustrative Example

The data in table 4 were obtained in an experiment designed to measure the absorption of gamma radiation by lead (Pb) and by aluminum (Al).

For each of the two metals, two sets of measurements were made, each of which involved 5 different values for the distance between the radioactive source and the counting instrument. The thickness of the absorbing metal was varied by stacking plates of identical thickness upon each other, and placing these stacks between the source and the counter. Five thicknesses were used, obtained by making stacks of respectively 1, 3, 6, 7, and 10 plates. The response is the logarithm, to the base 10, of the number of pulses per second. The thickness of a single plate was not the same for the two metals, but this does not invalidate our analysis.

TABLE 4. *Absorption of gamma radiation by lead and aluminum*[a]

Row, i	Distance, cm	Set	Number of plates, k_j: 1	3	6	7	10
			Lead				
1	3.8		1.801	1.765	1.696	1.670	1.606
2	5.2		1.621	1.572	1.516	1.486	1.425
3	6.0	I	1.526	1.481	1.406	1.401	1.333
4	9.0		1.222	1.169	1.102	1.078	1.010
5	12.5		0.973	0.939	0.862	0.850	0.781
6	3.8		1.805	1.768	1.704	1.680	1.615
7	5.2		1.609	1.572	1.511	1.482	1.408
8	6.0		1.494	1.461	1.387	1.324	1.315
9	9.0	II	1.233	1.208	1.130	1.111	1.046
10	12.5		0.978	0.930	0.870	0.844	0.779
			Aluminum				
11	3.8		1.834	1.818	1.811	1.790	1.777
12	5.2		1.632	1.613	1.600	1.603	1.597
13	6.0	I	1.509	1.482	1.476	1.454	1.447
14	9.0		1.249	1.224	1.204	1.211	1.179
15	12.5		0.986	0.971	0.966	0.960	0.943
16	3.8		1.916	1.913	1.884	1.887	1.871
17	5.2		1.732	1.723	1.698	1.696	1.674
18	6.0	II	1.632	1.624	1.592	1.588	1.579
19	9.0		1.344	1.341	1.312	1.311	1.290
20	12.5		1.118	1.118	1.106	1.086	1.066

[a] Tabulated value $= \log_{10} \frac{\text{pulses}}{\text{sec}}$ (corrected for background noise).

The first phase of the analysis was carried out on a two-way table, in which the rows represent combinations of three factors: metal, distance, and set. The columns of the table correspond to the 5 levels of "number of plates," k_j.

Using the notation introduced previously in this paper, the analysis of variance is given in table 5, where the M values given in table A1 were used for the degrees of freedom. The structural parameters are given in table 6.

TABLE 5. *Absorption of Gamma Radiation, Analysis of Variance* [a]

Source	*DF*	*SS*	*MS*	
Total	100	205.34167	.....	
μ	1	195.97757	.....	
ρ_i	19	9.106487	0.4793	
γ_j	4	0.190188	.0475	
η_{ij}	76	.067447	.000887	
$\theta u_i v_j$	31.9	.063801	.002000	
$\theta' u_i' v_j'$	21.3	.002083	.000098	
$\theta'' u_i'' v_j''$	14.3	.000847	.000059	0.000083
Residual [b]	8.5	.000716	.000084	

[a] See footnote a of table 3.

[b] The total number of multiplicative terms for these data is 4; hence the residual is identical with the fourth eigenvalue. The degrees of freedom for the residual are obtained by difference.

TABLE 6. *Absorption of Gamma Radiation, Structural Parameters*

Model: $z_{ij} = 1.40 + 1.3495 r_i + 0.0975 c_j + 0.2526 u_i v_j + \epsilon_{ij}$

Distance	Set	r_i	u_i	k_j	c_j	v_j
3.8		0.2280	0.2258	1	−0.6447	−0.6487
5.2		.0919	.2135			
6.0	Pb I	.0218	.2061	3	−.2494	−.2540
9.0		−.2102	.2623			
12.5		−.3845	.2148	6	−.0848	−.0767
3.8		.2330	.2043	7	.3556	.3689
5.2		.0863	.2336			
6.0	Pb II	−.0028	.2290	10	.6233	.6105
9.0		−.1884	.2117			
12.5		−.3851	.2225			
3.8		.3009	−.2117			
5.2		.1549	−.2850			
6.0	Al I	.0546	−.2009			
9.0		−.1382	−.1908			
12.5		−.3221	−.2652			
3.8		.3662	−.2359			
5.2		.2258	−.2036			
6.0	Al II	.1505	−.2052			
9.0		−.0595	−.2088			
12.5		−.2231	−.2166			

Since the mean squares in table 5 are small for the second, third, and fourth multiplicative terms, when compared to that of the first multiplicative term, the analysis of variance indicates quite conclusively the need for a single multiplicative term, in addition to the usual additive terms (grand average, row main effects and column main effects).

The model emerging from this analysis is given by the following equation:

$$z_{ij} = 1.400 + 1.3495 r_i + 0.0975 c_j + 0.2526 u_i v_j + \epsilon_{ij}$$

where r_i, u_i, c_j, and v_j are listed in table 6, and ϵ_{ij} is a random error, with standard deviation $\hat{\sigma} = 0.0091$.

This model may be simplified by examining the structural parameters. We first note the great similarity between v_j and c_j.[5]
Assuming that $v_j = c_j$, our model becomes

$$z_{ij} = 1.400 + 1.3495 r_i + (0.0975 + 0.2526 u_i) v_j + \epsilon_{ij} \tag{24}$$

Our next task is to study the structure of the parameters r_i and u_i in terms of the variables of which they are functions, i.e., distance and set. For the parameter u_i, a formal analysis is hardly necessary. It is evident from table 6 that u_i has essentially the same absolute value for all 20 combinations of distance with set, and that its sign is + for Pb and − for Al. Since $\sum_i u_i^2 = 1$. we obtain:

$$u_i = \pm \sqrt{\frac{1}{20}} = \pm 0.2236 \tag{25}$$

the + sign applying to Pb, and the − sign to Al.

To study r_i, we make a two-way analysis of the same type as above, considering r_i as the response, and identifying the rows with distance and the columns with the four sets (Pb I, Pb II, Al I, Al II). The results are shown in table 7a. From the analysis of variance it is clear that a simple additive model is appropriate. Thus, we write the model

$$r_i = A_d + B_s \tag{26}$$

with the values of the new parameters A_d and B_s listed in table 7b. Combining eqs (24), (25) and (26) yields:

$$z_{ij} = 1.400 + 1.3495 (A_d + B_s) + (0.0975 \pm 0.0565) v_j + \epsilon_{ij} \tag{27}$$

where the + sign applies to Pb and the − sign to Al.

TABLE 7a. *Analysis of variance of* r_i

Source	*DF*	*SS*	*MS*
Total	20	1.0000	0.05
Mean	1	0	0
Distance	4	0.9300	0.2325
Set	3	.0688	.0229
Distance × set	12	.001203	.000100
EV1 [a]	8.4	.000752	.000090
EV2	3.0	.000343	.000114
EV3	0.6	.000108	.000180

[a] EV1, EV2, and EV3 denote the three eigenvalues into which the interaction Distance × Set is partitoned.

[5] Here the advantage of "standardizing" the $\hat{\rho}_i$ and $\hat{\gamma}_j$ into r_i and c_j becomes apparent. An identical relationship between v_j and c_j (or between v_j and $(-c_j)$) corresponds to a proportional relationship between v_j and γ_j. Such a proportional relationship might have escaped attention, unless the values had been plotted on a graph.

TABLE 7b. *Values of parameters in equation:* $r_i = A_d + B_s$

Distance	A_d		Set	B_s
3.8	0.2820		Pb I	−0.0506
5.2	.1397		Pb II	−.0514
6.0	.0560		.....	.....
9.0	−.1491		Al I	.0100
12.5	−.3287		Al II	.0920

All that remains to be done is to fit appropriate curves to the parameter A_d, B_s, v_j as functions of the variables distance, metal, and k_j (number of plates) respectively, using the values listed in table 7b (for A_d and B_s) and in table 6 (for v_j).

In view of the near-identity of the two values of B_s for Pb, both sets for this metal can be expressed by the same equation. Equation (27) then leads to the following model equations:

For Pb (sets I and II):

$$z_{ij} = 1.331 + 1.3495A_d + 0.1540v_j + \epsilon_{ij}. \tag{28}$$

For Al:

set I: $z_{ij} = 1.414 + 1.3495A_d + 0.0410v_j + \epsilon_{ij}$ (29)

set II: $z_{ij} = 1.524 + 1.3495A_d + 0.0410v_j + \epsilon_{ij}$. (30)

A plot of A_d versus distance and of v_j versus number of plates would show smooth relationships for both these parameters. In fact, the plot of v_j versus the number of plates is simply a straight line, which is in accordance with the theory of absorption of radiation.

We need not complete this phase of the analysis, since it involves no novel features.

Factorial types of data, such as those used in our illustration (table 4) are often analyzed by the conventional analysis of variance, in which the sum of squares is partitioned into main effects and interactions. To interpret such an analysis, it must be ascertained whether the interactions that are found to be significant depend in a systematic way on the factors. When this is the case, the model is not additive, and then the nature of the interactions must be further elucidated. This is precisely what the method of analysis proposed in this paper is intended to do.

Thus, in our analysis of the illustrative example, in which the three factors "type of metal," "distance," and "sets" were merged into one (the "rows" of the table), a clear dichotomy was nevertheless indicated by the parameter u_i, which sharply differentiated between Pb and Al, and led to the entirely different coefficient of v_j in eq (28) as compared to eqs (29) and (30). This is of course an "interaction" between "type of metal" and "number of plates," but the advantage of our approach is that rather than merely indicating the *presence* of an interaction, the parameter leads to a quantitative expression for it.

8. Relationship to Models not Based on Principal Component Analysis

a. The Additive Model

Clearly, (23) expresses an additive model if and only if all multiplicative terms vanish (or become mere random experimental error). An additive model is therefore diagnosed if none of the mean squares obtained in the partitioning of the interaction are large with respect to subsequent mean squares.

b. The Linear Model

This model [7, 8] is given by the equation:

$$z_{ij} = \mu + \rho_i + \gamma_j + (\beta_i - 1)\gamma_j + \epsilon_{ij}$$

where $\bar{\beta}=1$.

It can be written as:

$$z_{ij}=\mu+Rr_i+Gc_j+(\beta_i-1)Gc_j+\epsilon_{ij}.$$

In order that this model apply, it is necessary that the following two conditions be fulfilled:

1. The partitioning of the interaction must yield only one significant product term;
2. We must have, either

$$v_j=c_j \qquad (\text{all } j)$$

or

$$v_j=-c_j \qquad (\text{all } j).$$

Obviously a linear model also holds when

$$z_{ij}=\mu+\rho_i+\gamma_j+\rho_i(\delta_j-1)+\epsilon_{ij}: \qquad \bar{\delta}=1.$$

In this case, condition 2 above becomes:

$$u_i=r_i \qquad (\text{all } i)$$

or

$$u_i=-r_i \qquad (\text{all } i)$$

c. The Concurrent Model

This model [7, 8] is a special case of b.

It is represented by the relation:

$$z_{ij}=\mu+\rho_i+\gamma_j+K\rho_i\gamma_j+\epsilon_{ij}.$$

A set of data will be represented by a concurrent model if and only if the following three conditions are simultaneously fulfilled:

1. Only one significant product term results from the partitioning of the interaction.
2. $u_i=r_i$ or $u_i=-r_i$ (all i)
3. $v_j=c_j$ or $v_j=-c_j$ (all j).

d. Tukey's One Degree of Freedom for Non-Additivity

In 1949, Tukey [13] proposed a test for nonadditivity consisting in the extraction of one degree of freedom from the row by column interaction. This test can be interpreted in terms of the model

$$z_{ij}=\mu+\rho_i+\gamma_j+K\rho_i\gamma_j+\epsilon_{ij}$$

which is identical with what we have called the concurrent model.

In Tukey's procedure, the sum of squares, with one degree of freedom, for $K\rho_i\gamma_j$, is tested against the remaining sum of squares, with $[(m-1)\ (n-1)-1]$ degrees of freedom.

We can regard this model as a special case of the linear model (case b).

If we write:

$$z_{ij}=\mu+\rho_i+\gamma_j+K\rho_i\gamma_j+[(\beta_i-1)-K\rho_i]\gamma_j+\epsilon_{ij}$$

the allocation of degrees of freedom is 1 for $K\rho_i\gamma_j$, and $(m-2)$ for $[(\beta_i-1)-K\rho_i]\gamma_j$, provided that β_i and K are estimated by the appropriate procedure [7]. Thus a valid test for $K\rho_i\gamma_j$ is obtained by computing the F ratio of the mean squares corresponding to these two terms, with 1 and $(m-2)$ degrees of freedom. Tukey's procedure consists in testing the term $K\rho_i\gamma_j$ versus the combined

mean square for $[(\beta_i-1)-K\rho_i]\gamma_j$ *and* ϵ_{ij}. The denominator now has $[(m-1)(n-1)-1]$ degrees of freedom. Thus carried out, the test overlooks the possible significance of the term $[(\beta_i-1)-K\rho_i]\gamma_j$. It appears preferable to consider first the linear model, and separate out one degree of freedom for concurrence. Indeed, Tukey's test will tend to lead to an erroneous conclusion of additivity everytime the model is linear without being concurrent.

e. The Vacuum-Cleaner Model

In a paper appearing in 1962, Tukey [14] presented what may be considered as a repeated twofold application of the linear model, both row-wise and column-wise. He referred to it as the "vacuum cleaner model." The first "sweep" of the vacuum cleaner model may be represented by the following equation:

$$z_{ij}=\mu+\rho_i+\gamma_j+K\rho_i\gamma_j+\beta_i\gamma_j+\rho_i\delta_j+\epsilon_{ij}$$

where

$$\sum_i \beta_i=\sum_j \delta_j=0$$

and

$$\sum_i \beta_i\rho_i=\sum_j \gamma_j\delta_j=0.$$

If data obeying this model are analyzed by the method of this paper, the first three eigenvalues will tend to extract all the information contained in the terms $K\rho_i\gamma_j$, $\beta_i\gamma_j$ and $\rho_i\delta_j$, but there will be no one-to-one-correspondence between the eigenvalues and these terms.

A characteristic feature of the vacuum cleaner model is that the "carriers" of each sweep are essentially the "coefficients" of the preceding sweep. For example, in the equation above, the three terms into which the interaction is partitioned are all linear functions of ρ_i and γ_j, the coefficients of the additive components. If an additional sweep were required, its term would be linear functions of β_i and δ_j, and so forth.

By contrast, the terms $\theta u_i v_j$ introduced consecutively in our method of analysis, are not made dependent upon each other. At each step, the *residuals alone* produce the new term $\theta u_i v_j$, whereas in the vacuum-cleaner model, the new terms are functions of the residuals *and* of the coefficients of the preceding sweep.

The preceding discussion shows that an analysis of two-way data carried out according to the method presented in this paper contains as special cases a number of methods not involving principal component analysis, and that application of this technique will allow us to recognize at once a number of frequently occurring special models.

9. Relationship to Other Approaches Involving Principal Component Analysis

As mentioned earlier in this paper, the idea of using principal component procedures in data analysis is not new. In the following we will show in what way the method presented in this paper differs from similar approaches.

The method of principal components originated in the field of psychology. We can visualize the basic problem in terms of a two-way table of m rows and n columns. Each row represents a different "subject" (e.g., a human being), and each column represents a particular psychological "test." The response is the score obtained by the ith "subject" in the jth "test." Here a vast conceptual difference exists between rows and columns. The object is to find a *small* number of linear combinations of the values in any particular row ("derived responses") that somehow typifies the overall response of the subject to the entire battery of tests. The model is formulated in terms of "characteristic vectors," equal in number to the "derived responses." For each derived response, the corresponding vector consists of the coefficients by which the n scores of any subject must be multiplied in order to obtain the linear combination yielding the derived response for that subject. The set of values obtained for all subjects for a particular derived response is denoted as the set of "scalar multipliers" corresponding to that vector.

In our model, no distinction in principle is made between vector components and scalar multipliers. They fulfill entirely analogous roles, one for the columns, and the other for the rows (they are the quantities denoted as v_j and u_i respectively). Furthermore, whereas in the classical approach, the calculation of principal components is generally carried out on the deviations of the original scores from the corresponding column means, we apply these calculations to the residuals from the additive model, i.e., after taking out both row and column means.

The literature dealing with the application of principal components to data analysis in the physical sciences contains some valuable contributions [1, 2, 3, 5, 6, 10, 12, 15, 17], but suffers largely from its adherence to the terminology and model building methods derived from the original field of application of this technique. In the first place, with the exception of reference [3], which will be discussed separately below, rows and columns are always treated as conceptually different entities. We have seen in our illustrative example that this distinction is unnecessary, even when the rows or columns represent discrete categories rather than controlled quantitative variables.

Secondly, one finds repeated reference to the "percentage of the variance accounted for by a particular latent root" [2, 5, 12]. While this may be a valid concept in psychological and similar applications, it could be a very misleading criterion in the analysis of physical and chemical data. Scientists in these fields generally base their acceptance of a residual sum of squares as an expression of experimental error on the agreement between the residual mean square and the variance of experimental error as known to them from previous experience. A good example is provided by Wernimont's data [15]. In analyzing jointly two spectrophotometers he obtains the two roots[6] (his Table III): $\lambda_1 = 7{,}896{,}525$ and $\lambda_2 = 2{,}553$, and a residual sum of squares of 151. The first root accounts for 99.9658 percent of the total sum of squares, and the second for only 0.0323 percent, yet Wernimont has shown that the second root is important and has given it a meaningful physical interpretation.

A third major difference between our approach and that of other authors (with the exception of [3]) is that we extract both components of the additive model (ρ_i and γ_j) *before* extracting characteristic roots. This leads, in the first place, to the possibility of treating rows and columns symmetrically, and, in the second place, to a set of completely "standardized" quantities (sum zero, sum of squares unity). As shown above, comparison of these quantities with the r_i and c_j allows for immediate recognition of important special cases.

The importance of this point can be seen from an examination of an illustrative example used by Simonds [12], and from his own analysis of these data. Starting with a 7×5 matrix of data, Simonds subjects the column average-corrected values to a principal component analysis, and decides that 2 vectors are sufficient to represent the data. His model is, accordingly:[7]

$$z_{ij} = \bar{z}_j + Y_{1,i} \cdot V_{1,j} + Y_{2,i} \cdot V_{2,j}.$$

The following values were obtained for the vectors $\bar{z}_j$ and $V_{1,j}$:

$$\bar{z}_j = [0.134 \quad 0.166 \quad 0.384 \quad 0.883 \quad 1.446]$$

$$V_{1,j} = [0.048612 \quad 0.066297 \quad 0.157030 \quad 0.401836 \quad 0.661341].$$

The vector $V_{1,j}$ is of course independent of the vector $\bar{z}_j$ since the latter had been removed from the data prior to the extraction of characteristic roots. Simonds fails to observe, however, that these two vectors are linearly related. If both vectors are normalized (by subtracting the mean and dividing by the square root of the sum of squares of deviation from the mean), one obtains:

$$\bar{z}_{j,\,\text{norm.}} = [-0.4195 \quad -0.3908 \quad -0.1957 \quad 0.2510 \quad 0.7550]$$

$$V_{1,j,\,\text{norm.}} = [\,-.4174 \quad -.3836 \quad -.2102 \quad .2577 \quad .7536].$$

[6] The first root given in Wernimont's table is actually 94,790,740. This value is in error; it should have been 7,896,525 [16].

[7] We follow as closely as possible his notation, modifying it only slightly, by introducing the subscripts i and j to represent rows and columns.

Thus, the model equation really contains one less vector than that given by Simonds. A finding of this type is of course important for the physical interpretation of data.

Similarly, Wernimont [15] takes no notice of the fact that, after normalization, the first characteristic vector he obtained in the analysis of single spectrophotometers, is identical with the vector of column averages, and that the corresponding vector of scalar multiples is identical with the vector of row averages. These facts, as seen above, indicate a concurrent model and this is precisely the model proposed by Wernimont. But an analysis in which both row and column averages have first been eliminated leads to a more exact testing procedure of the concurrence of the model.

Last, but not least, some of the pertinent papers that have come to the author's attention, including [3], suffer from a number of misconceptions regarding the proper number of degrees of freedom to be allocated to the successive eigenvalues. Others [2, 5, 12] ignore the matter altogether, by limiting their considerations to the "percentage of total variability explained" by the various eigenvalues. As pointed out earlier, the eigenvalues are genuine additive components of the interaction sum of squares, but they are *not* distributed as independent chi-square variates. Therefore the usual intuitive concepts concerning degrees of freedom do not apply. Our method is to define degrees of freedom as a quantity such that when the eigenvalue is divided by it, one obtains, for the case of random normal deviates of variance σ^2, an unbiased estimate of σ^2. In contrast, the following procedures are found in the literature. Morris and Morrissey [10], starting with a 42×31 matrix of original measurements, extract three roots, after initial subtraction of the column means. The residual sum of squares is then divided by 41×28, indicating that they regard as the proper formula for residual degrees of freedom, the expression $(m-1)\ (n-p)$, where p is the number of roots extracted. The formula seems to be based on the belief that the extraction of each root results in the loss of one degree of freedom for each of the $m-1$ independent rows. A similar formula is used by Wernimont [15]: here no adjustment (for the column-means) is made prior to the extraction of the roots, and the residual variance is calculated by dividing the residual sum of squares by $m(n-p)$. Judd et al. [6], appear to make no allowance, in terms of degrees of freedom for the fact that the eigenvalues are computed from the data. They state: "The variance for each set of data was computed in the usual way as the sum of squares of the differences between each input data and the corresponding value reconstituted from the mean and the first four characteristic vectors . . . divided by the number of input data." Gollob [3] defines mean squares corresponding to the partitioning of the interaction term on the basis of a heuristic argument. According to his definition, the "mean square" corresponding to the kth eigenvalue is the quotient of the eigenvalue by $(m+n-1-2k)$ (our notation).

A comparison was made between the values $m+n-1-2k$ and our empirical results. Table 8 lists both sets for matrices of various sizes. It is seen that the values $(m+n-1-2k)$ become less acceptable as m and n increase. In fact, whereas for a 4×4 matrix the value for $k=1$ is too low by about 21 percent, the corresponding value in a 16×16 matrix is too low by about 43 percent. The situation is reversed at the other end of the series of product terms (large k) where the estimate $(m+n-1-2k)$ is far too large. Thus, mean squares based on these "degrees of freedom" will be too large for the beginning terms and too small for the later terms. Comparisons of these "mean-squares" with an estimate of error obtained from within-cell replication will be vitiated by the biases inherent in these mean squares.

10. Computer Program

A computer program has been prepared by Mary N. Steel, for the application of the method of analysis presented in this paper, for data displayed in matrices of size up to 78×78. Details of this program, which is written in Fortran V, will be the subject of a separate publication.

The program includes, in addition to the analysis of variance for the additive effects, a partitioning of the row by column interaction into eigenvalues. It also tabulates the u and v vectors and the residuals after each successive step in the application of the method. Thus, residuals are given for the additive model, the model including one multiplicative term, two such terms, and so on.

TABLE 8. *Comparison of Gollob's formula for degrees of freedom with Monte-Carlo results of this study*

Term in partitioning of interaction	Size of matrix									
	4×4		6×6		8×8		12×12		16×16	
	G(1)	M(2)	G	M	G	M	G	M	G	M
1	5	6.45	9	13.35	13	20.72	21	35.39	29	51.08
2	3	1.97	7	6.96	11	12.37	19	25.28	27	39.68
3	1	.27	5	3.17	9	7.76	17	18.85	25	31.79
4			3	1.12	7	4.39	15	13.98	23	25.78
5			1	0.15	5	2.14	13	9.96	21	20.56
6					3	0.74	11	6.84	19	16.20
7					1	.11	9	4.46	17	12.59
8							7	2.62	15	9.56
9							5	1.31	13	7.00
10							3	0.48	11	4.86
11							1	.069	9	3.26
12									7	1.93
13									5	0.97
14									3	.34
15									1	.045

(1) Gollob's formula.
(2) Monte-Carlo results (this study).

All residuals are "normalized" through division by their root mean square, to facilitate the detection of outliers.

At this time, the degrees of freedom for the breakdown of the interaction term have been calculated for matrices of size up to 20×100. Thus, the calculation of mean squares in the partitioned interaction is also limited to matrices of this size.

The author wishes to express his gratitude to Mary N. Steel for performing the Monte-Carlo calculations and for developing the program for the application of this analysis.

11. Appendix

Tables A1, A2, and A3 were obtained by a Monte-Carlo experiment, as explained in the body of the paper. Each table consists of three parts, corresponding respectively to the largest, the second-largest, and the third-largest eigenvalue. The parts are identified by the numerals 1, 2, and 3 at the beginning of each row. The second numeral in each row label represents the value of m; the column labels represent the values of n. All three tables are of course symmetrical with respect to m and n.

For interpolation purposes, use table A3 rather than A1. For example, if M_1, for $m=11$ and $n=18$ is to be calculated, we obtain by interpolation from table A3:

$$\% M_1, \text{ for } m=11,\ n=18, = 26.31.$$

The total number of degrees of freedom being $(m-1)(n-1)$, we have:

$$M_1 = 26.29\% \text{ of } (10 \times 17),$$

or

$$M_1 = 44.7.$$

TABLE A1. *Expected values of eigenvalues for random normal deviates—*M *values*

m \ *n*	4	5	6	7	8	10	12	16	20
1–4	6.45	8.47	9.86	11.61	12.88	15.08	17.96	23.33	28.20
1–5	8.47	10.37	11.82	13.59	14.75	18.11	21.01	26.29	31.87
1–6	9.86	11.82	13.35	15.44	16.92	20.42	22.85	28.87	36.78
1–7	11.61	13.59	15.44	17.18	18.91	22.89	25.59	31.83	37.46
1–8	12.88	14.75	16.92	18.91	20.72	24.15	27.67	34.52	40.70
1–10	15.08	18.11	20.42	22.89	24.15	27.81	31.88	38.81	45.11
1–12	17.96	21.01	22.85	25.59	27.67	31.88	35.39	42.76	50.00
1–16	23.33	26.29	28.87	31.83	34.52	38.81	42.76	51.08	58.11
1–20	28.20	31.87	36.78	37.46	40.70	45.11	50.00	58.11	66.34
1–32	42.86	46.77	50.83	53.77	57.50	63.21	68.60	78.62	88.71
1–50	63.85	68.93	73.74	77.62	81.68	88.66	95.11	107.41	117.36
1–100	120.90	126.58	132.94	138.45	143.64	153.00	161.58	176.32	192.08
2–4	1.97	3.04	4.01	5.06	6.06	7.81	9.98	14.08	17.90
2–5	3.04	4.25	5.45	6.72	7.78	10.10	12.50	16.80	21.27
2–6	4.01	5.45	6.96	8.21	9.68	12.15	14.60	19.51	25.54
2–7	5.06	6.72	8.21	9.77	11.21	14.26	16.77	22.18	27.02
2–8	6.06	7.78	9.68	11.21	12.73	15.70	18.72	24.26	29.77
2–10	7.81	10.10	12.15	14.26	15.70	19.21	22.12	28.60	33.86
2–12	9.98	12.50	14.60	16.77	18.72	22.12	25.28	32.43	38.32
2–16	14.08	16.80	19.51	22.18	24.26	28.60	32.43	39.68	46.05
2–20	17.90	21.27	25.54	27.02	29.77	33.86	38.32	46.05	53.76
2–32	29.99	34.01	38.38	41.15	44.63	50.22	55.54	65.32	74.28
2–50	48.10	53.54	58.35	62.61	66.57	73.16	80.12	91.31	101.51
2–100	98.31	105.69	112.21	118.49	123.48	133.58	142.62	156.95	171.26
3–4	0.27	0.67	1.15	1.72	2.26	3.39	4.87	7.74	10.66
3–5	.67	1.38	2.22	3.08	3.72	5.42	7.11	10.65	14.34
3–6	1.15	2.22	3.17	4.15	5.23	7.28	9.15	13.18	18.06
3–7	1.72	3.08	4.15	5.44	6.48	8.95	11.05	15.55	19.69
3–8	2.26	3.72	5.23	6.48	7.76	10.39	12.80	17.61	22.35
3–10	3.39	5.42	7.28	8.95	10.39	13.38	16.03	21.33	26.30
3–12	4.87	7.11	9.15	11.05	12.80	16.03	18.85	25.03	30.50
3–16	7.74	10.65	13.18	15.55	17.61	21.33	25.03	31.79	37.73
3–20	10.66	14.34	18.06	19.69	22.35	26.30	30.50	37.73	45.04
3–32	20.14	25.11	29.34	32.34	35.65	41.24	46.31	55.84	64.41
3–50	35.12	41.62	47.57	51.70	55.72	62.49	68.76	80.16	90.24
3–100	79.50	89.40	97.12	103.02	109.26	119.49	127.74	143.20	156.38

TABLE A2. *Standard deviation of eigenvalues for random normal deviates*

m \ n	4	5	6	7	8	10	12	16	20
1–4	3.24	3.83	4.05	4.36	4.52	4.72	5.26	5.71	6.25
1–5	3.83	4.18	4.12	4.28	4.71	5.12	5.45	5.84	6.40
1–6	4.05	4.12	4.30	4.88	4.71	5.64	5.51	6.09	6.92
1–7	4.36	4.28	4.88	4.80	5.16	5.28	5.64	6.03	6.51
1–8	4.52	4.71	4.71	5.16	5.26	5.30	6.02	6.49	6.88
1–10	4.72	5.12	5.64	5.28	5.30	5.50	6.41	6.25	6.89
1–12	5.26	5.45	5.51	5.64	6.02	6.41	6.65	6.57	7.53
1–16	5.71	5.84	6.09	6.03	6.49	6.25	6.57	7.10	7.60
1–20	6.25	6.40	6.92	6.51	6.88	6.89	7.53	7.60	7.87
1–32	7.38	7.59	7.96	7.89	7.98	8.02	8.26	8.37	8.77
1–50	9.55	9.55	9.44	8.88	9.04	9.36	9.77	9.86	10.16
1–100	12.96	12.22	12.50	11.84	11.96	11.67	11.59	11.95	12.19
2–4	1.39	1.77	2.08	2.25	2.62	2.77	3.17	3.56	4.11
2–5	1.77	2.04	2.29	2.48	2.70	3.11	3.43	3.95	4.32
2–6	2.08	2.29	2.61	2.83	2.94	3.36	3.68	4.00	4.46
2–7	2.25	2.48	2.83	3.04	3.04	3.36	3.77	4.20	4.72
2–8	2.62	2.70	2.94	3.04	3.17	3.43	3.82	4.26	4.75
2–10	2.77	3.11	3.36	3.36	3.43	3.99	4.09	4.66	5.00
2–12	3.17	3.43	3.68	3.77	3.82	4.09	4.10	4.85	4.93
2–16	3.56	3.95	4.00	4.20	4.26	4.66	4.85	5.30	5.55
2–20	4.11	4.32	4.46	4.72	4.75	5.00	4.93	5.55	5.83
2–32	5.46	5.24	5.60	5.44	5.52	5.82	5.88	6.27	6.20
2–50	6.93	7.09	6.79	6.96	6.61	6.92	7.08	7.36	7.20
2–100	10.15	9.59	9.08	9.56	9.04	9.23	8.90	9.27	9.41
3–4	0.37	0.69	0.92	1.20	1.35	1.72	2.13	2.80	3.14
3–5	.69	.97	1.23	1.56	1.66	1.86	2.30	2.90	3.36
3–6	.92	1.23	1.53	1.75	2.00	2.21	2.46	3.08	3.44
3–7	1.20	1.56	1.75	2.00	2.07	2.43	2.66	3.32	3.62
3–8	1.35	1.66	2.00	2.07	2.31	2.62	2.80	3.19	3.56
3–10	1.72	1.86	2.21	2.43	2.62	2.95	3.11	3.53	3.68
3–12	2.13	2.30	2.46	2.66	2.80	3.11	3.34	3.60	3.87
3–16	2.80	2.90	3.08	3.32	3.19	3.53	3.60	4.17	4.45
3–20	3.14	3.36	3.44	3.62	3.56	3.68	3.87	4.45	4.76
3–32	4.59	4.29	4.53	4.55	4.60	4.52	4.98	5.09	5.38
3–50	6.54	5.64	5.68	5.67	5.84	5.78	5.77	5.87	6.01
3–100	9.21	8.27	8.12	7.77	7.98	7.83	7.36	8.11	8.05

TABLE A3. M *Expressed as percent of total interaction degrees of freedom*

m \ n	4	5	6	7	8	10	12	16	20
1–4	74.54	69.42	65.66	62.91	60.68	57.64	54.80	51.63	49.72
1–5	69.42	63.83	59.06	55.83	53.67	50.46	47.74	44.18	41.89
1–6	65.66	59.06	54.05	51.57	48.48	44.96	42.26	38.85	36.74
1–7	62.91	55.83	51.57	47.70	45.50	41.47	38.84	35.07	32.96
1–8	60.68	53.67	48.48	45.50	42.79	38.66	36.07	32.61	30.30
1–10	57.64	50.46	44.96	41.47	38.66	34.55	32.36	28.73	26.54
1–12	54.80	47.74	42.26	38.84	36.07	32.36	29.96	25.90	24.08
1–16	51.63	44.18	38.85	35.07	32.61	28.73	25.90	22.59	20.62
1–20	49.72	41.89	36.74	32.96	30.30	26.54	24.08	20.62	18.27
1–32	46.09	38.12	32.71	29.41	26.42	22.77	20.13	16.96	14.97
1–50	43.39	35.36	29.98	26.36	23.73	20.14	17.69	14.53	12.66
1–100	40.44	32.14	26.96	23.40	20.82	17.15	14.83	11.90	10.21
2–4	22.39	25.06	26.69	27.67	28.61	29.56	30.38	31.21	31.51
2–5	25.06	26.34	27.25	27.75	28.13	28.13	28.40	28.17	27.96
2–6	26.69	27.25	28.09	27.41	27.72	26.85	26.86	26.17	25.55
2–7	27.67	27.75	27.41	27.13	26.86	26.04	25.44	24.50	23.77
2–8	28.61	28.13	27.72	26.86	26.17	25.05	24.36	22.99	22.24
2–10	29.56	28.13	26.85	26.04	25.05	23.83	22.41	21.18	19.88
2–12	30.38	28.41	26.86	25.44	24.36	22.41	21.16	19.65	18.40
2–16	31.21	28.17	26.17	24.50	22.99	21.18	19.65	17.58	16.28
2–20	31.51	27.96	25.55	23.77	22.24	19.88	18.40	16.28	14.83
2–32	32.25	27.62	24.73	22.37	20.53	18.06	16.30	14.08	12.56
2–50	32.72	27.44	23.76	21.28	19.36	16.61	14.89	12.38	10.93
2–100	32.93	26.80	22.72	20.01	17.87	14.98	13.09	10.59	9.10
3–4	3.07	5.52	7.65	9.42	10.71	12.80	14.82	17.16	18.77
3–5	5.52	8.56	11.08	12.75	13.41	15.09	16.15	17.83	18.86
3–6	7.65	11.08	12.76	13.86	14.96	16.10	16.77	17.66	18.08
3–7	9.42	12.75	13.86	15.12	15.50	16.38	16.76	17.19	17.30
3–8	10.71	13.41	14.96	15.50	15.94	16.57	16.65	16.71	16.72
3–10	12.80	15.09	16.10	16.38	16.57	16.58	16.23	15.79	15.42
3–12	14.82	16.15	16.77	16.76	16.65	16.23	15.76	15.16	14.63
3–16	17.16	17.83	17.66	17.19	16.71	15.79	15.16	14.09	13.32
3–20	18.77	18.86	18.08	17.30	16.72	15.42	14.63	13.32	12.44
3–32	21.66	20.38	18.90	17.56	16.41	14.82	13.59	12.03	10.90
3–50	23.89	21.31	19.37	17.57	16.21	14.18	12.77	10.88	9.71
3–100	26.63	22.65	19.66	17.38	15.81	13.40	11.73	9.66	8.31

12. References

[1] Church, A., Analysis of data when the response is a curve, Technometrics **8,** 229–246 (1966).

[2] Condit, H. R., and Grum, F., Spectral energy distribution of daylight, J. Opt. Soc. Am. **54,** 937–944 (1964).

[3] Gollob, H. F., A Statistical model which combines features of factor analytic and analysis of variance techniques, Psychometrika **33,** 73–116 (1968).

[4] Hotelling, H., Analysis of a complex of statistical variables into principal components, Journal of Educational Psychology **24,** 417–441; 498–520 (1933).

[5] Jeffers, J. N. R., Two case studies in the application of principal component analysis, Applied Statistics, **16,** 225–236 (1967).

[6] Judd, D. B., MacAdam, D. L., and Wyszecki, G., Spectral distribution of typical daylight as a function of correlated color temperature, J. Opt. Soc. Am. **54,** 1031–1040 (1964).

[7] Mandel, J., Non-additivity in two-way analysis of variance, J. Am. Stat. Assoc., **56,** 878–888 (1961).

[8] Mandel, J., A method for fitting empirical surfaces to physical or chemical data, Technometrics **11,** 411–429 (1969).

[9] Mandel, J. and McCrackin, F. L., Analysis of families of curves, J. Res. Nat. Bur. Stand. (U.S.), 67A, (3) 259–267 (1969).

[10] Morris, R. H., and Morrissey, J. H., An objective method for determination of equivalent neutral densities of color film images, II. Determination of primary equipment, Society of America **44,** 530–534 (1954).

[11] Pike, E. W., and Silverberg, T. R., Designing mechanical computers, Machine Design **24,** 131–137; 159–163 (1952).

[12] Simonds, J. L., Application of characteristic vector analysis to photographic and optical response data, J. Opt. Soc. Am. **53,** 968–974 (1963).

[13] Tukey, J. W., One degree of freedom for non-additivity, Biometrics **5,** 232–242 (1949).

[14] Tukey, J. W., The future of data analysis, Anns. Math. Stat. **33,** 1–67 (1962).

[15] Wernimont, G., Evaluating laboratory performance of spectrophotometers, Anal. Chem. **39,** 554–562 (1967).

[16] Wernimont, G., private communication.

[17] Williams, E. J., The interpretation of interactions in factorial experiments, Biometrika **39,** 65–81 (1952).

(Paper 73B4–311)