PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

THE EVALUATION OF MULTIPLE AND PARTIAL CORRELATION COEFFICIENTS FROM THE FACTORIAL MATRIX*

P. S. DWYER
THE UNIVERSITY OF MICHIGAN

This paper shows how to compute multiple correlation coefficients, partial correlation coefficients, and regression coefficients from the factorial matrix. Special emphasis is given to computation technique and to approximation formulas. The method is extremely flexible in application since it may be applied to any subset of the original set of observed variables. It is also extremely useful when many of these coefficients are desired.

PART I

MULTIPLE CORRELATION COEFFICIENTS FROM THE FACTORIAL MATRIX

1. *Introduction.* One of the disadvantages of multiple correlation theory is the enormous amount of time needed to compute the multiple correlations when n is large. Even the Doolittle method, **(1, 2, 3)** which is being recognized as an efficient method of computing multiple correlations, demands a large amount of computation when n is large. It is shown in this part how multiple correlations can be obtained from the factorial matrix, and in particular, from a centroid solution, though any solution which gives communality and uniqueness factors is satisfactory. Though a centroid solution does take time, the total time necessary to obtain the multiple correlations is but a small fraction of that demanded by other methods if the number of variables is large, if the number of factors is small, and if a number of such coefficients is desired.

2. *Notation.* (a) *From multiple correlation theory.* We assume that there are n standard variables, $x_1, x_2, \cdots, x_i, \cdots, x_j, \cdots, x_n$.

**Editor's Note*: The reader will recognize the subject-matter treated in this article as closely allied to that of an article by Louis Guttman, "Multiple Rectilinear Prediction and the Resolution into Components," in the June, 1940 issue of this journal. Although there is considerable overlapping in the topics considered in these two articles, it was felt that some readers would be more interested in one approach, while others would gain more from the other approach and that still others would find both presentations of value. Guttman's article is somewhat more concerned with the theoretical aspects of the problem, while Dwyer's article emphasizes the technique of computation and approximation. The manuscript for Dwyer's article was received while Guttman's article was in press. At Dwyer's request, Guttman made available to him a pre-publication copy of his manuscript so that Dwyer was able to insert in his manuscript references to the earlier one.

The variables are recorded for a population of N individuals and the matrix of the correlations, $[r_{ij}]$, is computed The determinant of this matrix is indicated by Δ and the determinant of the matrix formed by deleting variable x_j is denoted by Δ_j. If R_j indicates the multiple correlation of x_j with the $n - 1$ remaining variables, then by the usual formula (4, p. 301)

$$R_j = \sqrt{1 - \frac{\Delta}{\Delta_j}}. \tag{1}$$

$R_{j)i(}$ is used to indicate the multiple correlation of x_j with all the remaining variables except x_i. Similarly $R_{j)ik(}$ indicates the multiple correlation of x_j with all the remaining variables except x_i and x_k.

(b). *From multiple factor theory.* We assume that the correlation matrix has been subjected to an analysis which results in $r + n$ standard orthogonal variables, r of which are common variables and n of which are uniqueness variables. More specifically we assume that every x_j can be represented by

$$x_j = a_{j1}y_1 + a_{j2}y_2 + \cdots + a_{jk}y_k + \cdots + a_{jr}y_r + u_jz_j, \tag{2}$$

where $y_1, y_2, \cdots, y_r, z_1, z_2, \cdots, z_n$ are the orthogonal variables, and the values a_{jk} are known.

It follows at once, since $r_{ij} = \frac{\sum x_i x_j}{N}$, that

$$r_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{ik}a_{jk} + \cdots + a_{ir}a_{jr}, \quad i \neq j \tag{3}$$

and

$$1 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jk}^2 + \cdots + a_{jr}^2 + u_j^2, \tag{4}$$

as indicated by Thurstone (5, p. 66). Furthermore we define $h_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jr}^2$ to be the communality of variable j and u_j^2 to be the "uniqueness." Then

$$h_j^2 + u_j^2 = 1 \quad \text{and} \quad u_j = \sqrt{1 - h_j^2}. \tag{5}$$

We next define the common part of x_j to be that part of x_j which is expressible in terms of the common variables y_k. Then

$$x'_j = a_{j1}y_1 + a_{j2}y_2 + \cdots + a_{jk}y_k + \cdots + a_{jr}y_r.$$

It is to be noted that x'_j is not a standard variable since its variance is not unity. However, if every factor loading a_{jk} is divided by h_j we get the standard variable

$$x''_j = \sum_{k=1}^{r} A_{jk}y_k,$$

where $A_{jk} = \frac{a_{jk}}{h_j}$. The factor loadings are increased by this process and hence we use the term "augmented" (5, p. 158). Furthermore, we introduce the term "communality augmented" to distinguish these from the "uniqueness augmented" loadings introduced in this paper.

We let $U^2 = [u_i^2]$ indicate the diagonal matrix composed of the uniqueness loadings. Then (3), (4), and (5) can be combined to give the fundamental matrix equation

$$R = AA' + U^2 . \tag{6}$$

3. *The value of* R_j. R_j can be expressed in terms of the factor loadings if Δ and Δ_j can be so expressed. Now Δ_j can be so expressed if Δ can. Hence such an expression for Δ will, essentially, solve our problem. A general development using matrix terminology is preceded by a special case, $n = 3$ and $r = 2$, in which determinant theory is used. We first express Δ in the form

$$\Delta = \begin{vmatrix} 1 & r_{12} & r_{13} & a_{11} & a_{12} \\ r_{21} & 1 & r_{23} & a_{21} & a_{22} \\ r_{31} & r_{32} & 1 & a_{31} & a_{32} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} \tag{7}$$

and eliminate the correlation coefficients by multiplying column 4 by $-a_{11}$, column 5 by $-a_{12}$, and adding to column 1; by multiplying column 4 by $-a_{21}$, column 5 by $-a_{22}$, and adding to column 2; by multiplying column 4 by $-a_{31}$, column 5 by $-a_{32}$, and adding to column 3. We then have

$$\Delta = \begin{vmatrix} u_1^2 & 0 & 0 & a_{11} & a_{12} \\ 0 & u_2^2 & 0 & a_{21} & a_{22} \\ 0 & 0 & u_3^2 & a_{31} & a_{32} \\ -a_{11} & -a_{21} & -a_{31} & 1 & 0 \\ -a_{12} & -a_{22} & -a_{32} & 0 & 1 \end{vmatrix} \tag{8}$$

We now divide the first row and the first column by u_1, the second row and the second column by u_2, the third row and the third column by u_3, and we get, if we let $\frac{a_{jk}}{u_j} = B_{jk}$, the "uniqueness augmented" loading,

$$\underline{\Delta} = u_1^2 u_2^2 u_3^2 \begin{vmatrix} 1 & 0 & 0 & B_{11} & B_{12} \\ 0 & 1 & 0 & B_{21} & B_{22} \\ 0 & 0 & 1 & B_{31} & B_{32} \\ -B_{11} & -B_{21} & -B_{31} & 1 & 0 \\ -B_{12} & -B_{22} & -B_{32} & 0 & 1 \end{vmatrix} \tag{9}$$

This can be further simplified by multiplying the first row by B_{11}, the second by B_{21}, the third by B_{31} and adding to the fourth row; by multiplying the first row by B_{12}, the second by B_{22}, the third by B_{32} and adding to the fifth row, so that we have

$$\Delta = u_1^2 u_2^2 u_3^2 \begin{vmatrix} 1+\sum B_{i1}^2 & \sum B_{i1}B_{i2} \\ \sum B_{i1}B_{i2} & 1+\sum B_{i2}^2 \end{vmatrix} . \tag{10}$$

This can be written as

$$\Delta = u_1^2 u_2^2 u_3^2 \begin{vmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{vmatrix} . \tag{11}$$

In the general case we have

$$\Delta[R] = \Delta \begin{bmatrix} RA \\ 0\ I \end{bmatrix} = \Delta \begin{bmatrix} R - AA'A \\ -A' \quad I \end{bmatrix} = \Delta \begin{bmatrix} U^2 A \\ -A'\ I \end{bmatrix} = \Delta[u^2]\Delta \begin{bmatrix} I\ B \\ -B'\ I \end{bmatrix}$$

$$= \Delta\,[u^2]\Delta \begin{bmatrix} I & B \\ 0\ I+BB' \end{bmatrix} = \Delta\,[u^2]\ \Delta\,[I+BB'] . \tag{12}$$

The determinant, $\Delta[I + BB']$, is a determinant of r rows and r columns, the elements of which are

$$M_{kl} = \begin{cases} \sum_{i=1}^{n} B_{ik}B_{il} & \text{if} \quad k \neq l \\ 1+\sum_{i=1}^{n} B_{ik}B_{il} & \text{if} \quad k = l \end{cases} . \tag{13}$$

The matrix $[I + BB']$ is the equivalent of Guttman's matrix Q (12, section 4). We may use the above result to indicate the value of Δ_j if variable j is deleted. Then (1) becomes

$$R_j = \sqrt{1 - u_j^2 \frac{\Delta[I+BB']}{\Delta_j[I+BB']}} . \tag{14}$$

where $\Delta_j(I + BB')$ is a determinant having r rows and r columns whose elements are given by

$$M_{kl)j(} = \left| \begin{array}{ll} \sum_{i=1}^{n} B_{ik}B_{il} - B_{jk}B_{jl} & \text{if} \quad k \neq l \\ 1 + \sum_{i=1}^{n} B_{ik}B_{il} - B_{jk}B_{jl} & \text{if} \quad k = l \end{array} \right| . \tag{15}$$

See also (12, Section 4).

The formula (14) is easy to apply (particularly if r is small) as each element of Δ_j can be obtained from the corresponding element of Δ by a simple subtraction of a product. The flexibility of the formula is also apparent since the correlation of any variable with any subset of the original n variables is easily effected. It is only necessary to complete the values of Δ and Δ_j from the factor loadings of the variables included.

4. *Illustration.* As a first example we use an illustration of Thurstone (5, p. 128) in which the six-place residuals are 0. In the factorial solution the correlation matrix involving 7 variables is replaced by a factorial matrix of two factors. The factor loadings, communality, and uniqueness for each variable are given in columns (2), (3), (4), and (5), while the "uniqueness augmented" loadings are given in columns (7) and (8).

TABLE I

Computation of "Uniqueness Augmented" Factor Loadings

(1) Variable j	(2) a_{j1}	(3) a_{j2}	(4) h_j^2	(5) u_j^2	(6) u_j	(7) B_{j1}	(8) B_{j2}
1	.659828	.120945	.450000	.550000	.741620	.889712	.163082
2	.830332	.265611	.760000	.240000	.489898	1.694908	.542176
3	—.541290	.637969	.700000	.300000	.547723	—.988255	1.164766
4	—.126124	.770774	.610000	.390000	.624500	—.201960	1.234226
5	.437356	.590526	.540000	.460000	.678233	.644846	.870683
6	.637638	—.336776	.520000	.480000	.692820	.920351	—.486094
7	.904489	.109084	.830000	.170000	.412311	2.193706	.264567

In getting the value R_1, we find that

$$M_{11} = 11.756955,\ M_{12} = M_{21} = .358146,\ M_{22} = 5.264916,$$

$$M_{11)1(} = 10.965368,\ M_{12)1(} = M_{21)1(} = .213050,$$

$$M_{22)1(} = 5.238320,$$

so that $\Delta[1 + BB'] = 61.771112$, $\Delta_1[I + BB'] = 57.394716$,

$$\frac{\Delta[I + BB']}{\Delta_1[I + BB']} = 1.076251,$$

and

$$R_1 = \sqrt{1 - (.550000)\,(1.076251)} = .638797.$$

This multiple correlation coefficient would be tedious to compute by conventional methods, so, in order to provide a check which the reader can follow with little computation, I select the case in which x_1 is predicted from x_2 and x_3. In this case

$M_{11} = 5.640949$, $M_{12} = M_{21} = -.087051$, $M_{22} = 2.677230$,
$M_{11)1(} = 4.849361$, $M_{12)1(} = M_{21)1(} = -.232147$, $M_{22)1(} = 2.650635$,
with $\Delta[I + BB'] = 15.094540$, $\Delta_1[I + BB'] = 12.799994$, and

$$R_1(= R_{1.23}) = \sqrt{1 - (.550000)(1.179261)} = .592795.$$

If we use the formula $R_1 = \sqrt{r_{12}\frac{\Delta_{12}}{\Delta_{11}} - r_{13}\frac{\Delta_{13}}{\Delta_{11}}}$, we get

$$R_1 = \sqrt{\frac{(.58)(.5016) + (.28)(.1176)}{.9216}} = .592795.$$

The conventional formulas are preferable for the case of three variables, but the advantage of the new method increases as the number of variables increases. For example, it was possible to obtain the seven multiple correlations obtained by estimating each variable from all the others from the three-place factor loadings in about three-quarters of an hour. The work is outlined in Table II.

There are $n(2^{n-1} - n)$ multiple correlation coefficients. Any one of these can be obtained from the uniqueness augmented factor loadings. Thus in the illustration there are 399 different multiple correlations, any one of which can be obtained in a relatively short time.

5. *Approximate factorial solutions.* The problem above is one in which the residuals are 0, to the required degree of approximation, and the multiple correlations agree to the desired number of places with the multiple correlations obtained by classical methods. Of course, if the multiple factor process is stopped before the residuals are actually zero—if the matrix equation $R = AA' + U^2$ is only approximately satisfied—the application of this method will give but approximations to the actual multiple correlations and the degree of the approximation will depend upon the smallness of the residuals.

Suppose for example a factor analysis of three variables resulted in the same factor loadings as those of variables 1, 2, 3, in Table II, but with second-order residuals $r_{12.2} = -.03$; $r_{13.2} = .04$; $r_{23.2} = .02$. Then the multiple correlation $R_{1.23}$, as indicated by this method, would be .593. The actual value of $R_{1.23}$, as determined by the correlations $r_{12} = .55$, $r_{13} = -.24$, $r_{23} = -.26$ is .559 so that the approximate solution has an error whose absolute value is .034. The actual average of the absolute values of the residuals is .030. If the residuals had

TABLE II

R_j for $j = 1, 2, 3, 4, 5, 6, 7$.

(1) Var.	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
j	a_{j1}	a_{j2}	h_j^2	u_j^2	u_j	B_{j1}	B_{j2}	$M_{11)j(}$	$M_{12)j(}$	$M_{22)j(}$	$\Delta_j[I+BB']$	$\frac{\Delta[I+BB']}{\Delta_j[I+BB']}$	R_j^2	R_j
1	.660	.121	.450	.550	.742	.889	.163	10.963	.218	5.239	57.387	1.0762	.408	.639
2	.830	.266	.760	.240	.490	1.694	.543	8.883	—.557	4.971	43.847	1.4085	.662	.814
3	—.541	.638	.700	.300	.548	—.987	1.164	10.779	1.512	3.911	39.871	1.5490	.535	.731
4	—.126	.771	.610	.390	.625	—.202	1.234	11.712	.612	3.743	43.463	1.4210	.446	.668
5	.437	.591	.540	.460	.678	.645	.872	11.337	—.200	4.505	51.033	1.2102	.443	.666
6	.638	—.337	.520	.480	.693	.921	—.486	10.905	.810	5.030	54.196	1.1396	.453	.673
7	.904	.109	.830	.170	.412	2.194	.265	6.939	—.218	5.196	36.008	1.7152	.708	.841
								11.753 M_{11}	.363 M_{12}	5.266 M_{22}	61.760 $\Delta[I+BB']$			

been $r_{12.2} = -.003$, $r_{13.2} = .004$, and $r_{23.2} = .002$, with an average absolute value of .003, the resulting error in $R_{1.23}$ turns out to have an absolute value of .0034. This subject needs further investigation, but it appears that there may be some definite relation between the size of residuals and the size of the error in the coefficient. This question might well be added to those recently raised by Wilson and Worcester (6, pp. 137-140) on the adequacy of reproduction of correlation coefficients.

6. *The case of a single factor.* In the case of a single factor the formula (14) becomes

$$R_j = \sqrt{1 - u_j^2 \frac{(1 + \sum_{i=1}^{n} B_{i1}^2)}{1 + \sum_{i=1}^{n} B_{i1}^2 - B_{j1}^2}}. \qquad (16)$$

As an illustration, we take a one-factor problem involving eight variables which Thurstone (5, p. 148) reproduced from Holzinger. Thurstone gave the factor loadings a_{j1}. These are given in Table III, and the values R_j are computed from them.

TABLE III

Multiple Correlations in the Case of a Single Factor

(1) Var. j	(2) a_{j1}	(3) h_j^2	(4) u_j^2	(5) B_{j1}^2	(6) $M_{11)j(}$	(7) $\frac{M_{11}}{M_{11)j(}}$	(8) K_j^2	(9) R_j^2	(10) R_j
1	.765	.586	.414	1.415	6.343	1.223	.506	.494	.703
2	.739	.546	.454	1.203	6.555	1.184	.538	.462	.680
3	.716	.512	.488	1.049	6.709	1.156	.564	.436	.660
4	.672	.451	.549	.821	6.937	1.118	.614	.386	.621
5	.634	.402	.598	.672	7.086	1.095	.655	.345	.587
6	.597	.357	.643	.555	7.203	1.077	.693	.307	.554
7	.595	.354	.646	.548	7.210	1.076	.695	.305	.552
8	.576	.331	.669	.495	7.263	1.068	.714	.286	.535
					$7.758 = M_{11}$				

These results are but approximations to the actual multiple correlations, since the first-order residuals (5, p. 148) are not all zero.

The multiple correlation formula may appear in different forms when there is but a single factor. Using $u_j^2 = 1 - h_j^2$ and

$$B_{j1}^2 = \frac{a_{j1}^2}{u_j^2} = \frac{h_j^2}{u_j^2},$$

(16) reduces to

$$R_j = h_j \sqrt{1 - \frac{1}{1 + \sum_{i=1}^{n} B_{i1}^2 - B_{j1}^2}}, \tag{17}$$

and it follows at once that $R_j < h_j$ so that h_j serves as an upper bound for R_j. In general, too, as the number of variables increases without limit the value $1 + \sum B_{i1}^2 - B_{j1}^2$ increases without limit so that R_j converges to h_j.

This convergence of R_j to h_j has been surmised by Roff (7, p. 6) who demonstrated it for the special case in which the r_{j1} are all equal and the r_{jk}, $k \neq 1$, are also all equal. His result can be obtained from (17) by utilizing the fact that $a_{21} = a_{31} = \cdots = a_{n1}$, $a_{j1}^2 = r_{nk}$ when $j \neq 1$; $a_{11}a_{j1} = r_{j1}$ so that $a_{j1} = \sqrt{r_{jk}}$ and $a_{11} = \dfrac{r_{j1}}{\sqrt{r_{ik}}}$. We have then

$$R_1 = r_{j1} \sqrt{\frac{n-1}{1 + (n-2)r_{jk}}}, \tag{18}$$

which is the formula given by Roff (7, p. 6) if n is replaced by $n + 1$.

7. *The case of many factors. Approximation method.* An approximation method is desired in case there are many factors, as the evaluation of the rth-order determinants demands considerable work. The approximation method is based on the tendency, when the number of variables is large, for the non-diagonal terms of $[I + BB']$ to be small in comparison with the diagonal terms. This situation has been used previously, (8, p. 175) in determining approximate values of the factor loadings of additional variables and in obtaining approximations to the roots of the characteristic equation. In the previous case it was assumed that $\sum a_{ik}a_{il} = 0$, while now it is assumed that $\sum B_{ik}B_{il} = M_{kl} = 0$ if $k \neq l$.

This condition is approximately satisfied by many problems even when the number of variables is relatively small. For example, the entries in column (10) of Table II are small compared with the entries in columns (9) and (11). If these terms are neglected, we get columns (12), (13), (14), and (15) of Table IV. Column (12) was obtained by multiplying the entries in columns (9) and (11) of Table II. Column (16), the correct correlation, is inserted to facilitate comparison with column (15), the approximate correlation obtained by neglecting non-diagonal terms.

The use of such approximation methods simplifies the work greatly if there are more than three factors. The formula becomes

$$R_j = \sqrt{1 - u_j^2 \left(\frac{M_{11}}{M_{11)j(}}\right)\left(\frac{M_{22}}{M_{22)j(}}\right) \cdots \left(\frac{M_{rr}}{M_{rr)j(}}\right)}. \tag{19}$$

TABLE IV

Approximate Solution of First Problem

Variable	(12) $\frac{\Delta_j[I+BB']}{}$	(13) $\frac{\Delta[I+BB']}{\Delta_j[I+BB']}$	(14) R_j^2	(15) R_j	(16) Actual R_j
1	57.435	1.0776	.407	.638	.639
2	44.157	1.0416	.750	.866	.814
3	42.157	1.4681	.560	.748	.731
4	*43.838*	(1.4118)	.449	.670	.668
5	51.073	1.2118	.443	.666	.666
6	54.852	1.1283	.458	.677	.673
7	36.055	1.7166	.708	.841	.841
	$\Delta = 61.891$				

In general, as the number of variables increases, (19) gives better approximations. Also $\frac{M_{11}}{M_{11)j(}} > 1$ and the largest possible value of R_j is $\sqrt{1 - u_j^2} = h_j$. Also, if periodically there is at least one $B_j > \varepsilon$, $\frac{M_{ii}}{M_{ii)j(}} \to 1$ as $n \to \infty$ so that $R_j \to h_j$. See also (12, Section 14).

In a previous study (9, p. 10), the square root of the communality was used as an upper bound for the multiple correlation coefficient. It is possible, with the use of (19), to obtain a good estimate of the correlation itself. In Table V are presented the values h_j used previously and the new values R_j obtained with (19).

TABLE V

Values of R_j and h_j for 19 Variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
h_j	.86	.91	.92	.94	.98	.92	.93	.95	.78	.93	.93	.90	.71	.90	.58	.89	.78	.79	.91
R_j	.85	.88	.88	.91	.95	.90	.91	.93	.77	.91	.90	.89	.69	.88	.56	.87	.77	.76	.89

8. *Use of non-uniqueness methods.* The development of the above formulas demands uniqueness factors since u_j must be different from 0. This appears to be a serious restriction at first, and yet it is not so serious in practice, since observed variables usually have unique components (5, p. 130). However, it does not appear at first that methods which assume no uniqueness factor (methods which use $AA' = R$ as the basic matrix equation) should yield results which are adaptable to the formulas of this paper. Of course, if the attempt is made to distribute the variance completely among r common components, *and the attempt is successful,* then the u_j are zero and the formulas above are not applicable. However, this attempt is frequently not entirely successful, particularly if the number of observed vari-

ables is large, and that is the case where the method is especially valuable. In an unsuccessful attempt to distribute the variance among r factors, the residuals of the diagonal terms are appreciably larger than the residuals of the non-diagonal terms. Hence, the sums of the squares of the factor loadings are appreciably less than unity when the sums of the products of the factor loadings do not differ appreciably from the correlations. In such a case, even though the basic assumption is that $u_j = 0$, the residuals of the diagonal terms may be used as the u_j^2 and the multiple correlation coefficients may be computed by the methods of this paper. The case of perfect prediction, where the u_j are actually as well as theoretically zero, has been discussed by Guttman (12, section 17).

PART II

Partial Correlations and Regression Coefficients from the Factorial Matrix

1. *Introduction.* It has been shown in Part I how the multiple correlation of an observed variable with other observed variables can be obtained from the factorial matrix. It is the purpose of this part to show how partial correlation coefficients and regression coefficients can be obtained similarly. The general plan is similar to that of Part I, though the matrices involved are not so simple. We use the well-known determinantal formulas

$$r_{12\cdot 34\cdots n} = \frac{\Delta_{12}}{\sqrt{\Delta_1 \Delta_2}}, \tag{1}$$

$$\beta_{12\cdot 34\cdots n} = \frac{\Delta_{12}}{\Delta_1}, \tag{2}$$

$$\beta_{21\cdot 34\cdots n} = \frac{\Delta_{21}}{\Delta_2}. \tag{3}$$

It has been shown, (Part I, Equation 12) that

$$\Delta_i = u^2_1 u^2_2 \cdots u^2_{i-1} u^2_{i+1} \cdots u^2_n \Delta_i[I + BB'] \tag{4}$$

so that Δ_1 and Δ_2 can be evaluated from the factorial matrix. The main task then is to evaluate Δ_{12}.

2. *The evaluation of* Δ_{12}. Perhaps the simplest approach is to carry through an evaluation of Δ in which, for a number of steps, the value of the minor is not changed. Thus if $\Delta_{12}[R]$ represents the determinant of the correlation matrix with the first row and second column eliminated, then

$$\Delta_{12}[R] = \Delta_{12}\begin{bmatrix} R & A \\ O & I \end{bmatrix} = \Delta_{12}\begin{bmatrix} R - AA' & A \\ -A' & I \end{bmatrix} = \Delta_{12}\begin{bmatrix} U^2 & A \\ -A' & I \end{bmatrix}.$$

We now divide the first column of $\Delta_{12}\begin{bmatrix} U^2 & A \\ -A' & I \end{bmatrix}$ by u_2, the first row by u_1, and the $n - 2$ following rows and columns by $u_2, \cdots, u_n$, respectively. The result can be written

$$\Delta_{12} = u_1u_2u_3^2u_4^2 \cdots u_n^2 \Delta \begin{bmatrix} 0 & 0 & B_2 \\ 0 & I & B_{)12(} \\ -B'_1 & -B'_{)12(} & I \end{bmatrix},$$

where B_2 is a single-rowed matrix consisting of $\frac{a_{21}}{u_2}, \frac{a_{22}}{u_2}, \frac{a_{23}}{u_2}, \cdots, \frac{a_{2r}}{u_2}$;

B_1 is a single-column matrix composed of elements $\frac{a_{1i}}{u_i}$; $B_{)12(}$ is a matrix of $n - 2$ rows and r columns in which no one of the first subscripts is 1 or 2. $B'_{)12(}$ is the transpose of $B_{)12(}$. It follows then that

$$\begin{aligned} \Delta_{12} &= u_1u_2u_3^2 \cdots u_n^2 \Delta \begin{bmatrix} 0 & 0 & B_2 \\ 0 & I & B_{)12(} \\ -B'_1 & 0 & I + B_{)12(} \quad B'_{)12(} \end{bmatrix} \\ &= u_1u_2u_3^2 \cdots u_n^2 \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + B_{)12(} \quad B'_{)12(} \end{bmatrix} \\ &= u_1u_2u_3^2 \cdots u_n^2 \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + BB' \end{bmatrix}. \end{aligned} \tag{5}$$

See also (12, section 5).

3. *The formula for partial correlation.* It follows, by substitution of (4) and (5) in (1) that

$$r_{12\cdot34\cdots n} = \frac{\Delta_{12}}{\sqrt{\Delta_1\Delta_2}} = \frac{u_1u_2u_3^2 \cdots u_n^2 \Delta \begin{bmatrix} 0 & B_2 \\ -B_1' & I + BB' \end{bmatrix}}{\sqrt{[u_2^2u_3^2 \cdots u_n^2 \Delta_1(I + BB')][u_1^2u_3^2 \cdots u_n^2 \Delta_2(I + BB')]}}$$

$$= \frac{\Delta \begin{bmatrix} 0 & B_2 \\ -B_1' & I + BB' \end{bmatrix}}{\sqrt{\Delta_1[I + BB'] \; \Delta_2[I + BB']}}. \tag{6}$$

The determinant in the numerator has $r + 1$ rows and columns, while the determinants in the denominator have but r rows and columns. The matrix $[I + BB']$ consists of r rows and r columns having elements

$$M_{kl} = \begin{cases} \sum B_{ik} B_{il} & \text{when } k \neq l \\ 1 + \sum B_{ik}\ B_{il} & \text{when } k = l \end{cases} \text{ with } B_{ik} = \frac{a_{ik}}{u_i}.$$

The determinants of the denominator are similar to the determinant of $[I + BB']$, but the variable x_1 is omitted in the first case and the variable x_2 in the second. The calculation is relatively simple and follows quickly from results used in getting the multiple correlation.

It is readily established that, more generally

$$r_{ij\cdot 12\cdots n} = \frac{(-1)^{i+j+1} \Delta \begin{bmatrix} 0 & B_j \\ -B_i' & I + BB' \end{bmatrix}}{\sqrt{\Delta_i [I + BB']\ \Delta_j [I + BB']}}. \tag{7}$$

The reader should compare (12, section 7).

4. *Illustration.* As an illustration we take the two-factor problem used as an example in the case of multiple correlation. We take out the first three variables, for ease in checking by classical methods, and compute $r_{12\cdot 3}$. The factor loadings are augmented to form the B's and the values M_{kl} determined. We find, from Table I of Part I, $B_{11} = .889712$, $B_{12} = .163082$, $B_{21} = 1.694908$, $B_{22} = .542176$,

$$M_{11} = 5.640949, \quad M_{12} = M_{21} = -.087051, \quad M_{22} = 2.677230,$$

with

$$\Delta_1 [I + BB'] = 12.799994, \quad \Delta_2 [I + BB'] = 5.585452.$$

$$\Delta \begin{bmatrix} 0 & B_2 \\ -B_i' & I + BB' \end{bmatrix} = \begin{vmatrix} .000000 & 1.694908 & .542618 \\ -.889712 & 5.640949 & -.087051 \\ -.163082 & -.087051 & 2.677230 \end{vmatrix} = 4.602031$$

so that

$$r_{12\cdot 3} = \frac{4.602031}{\sqrt{(12.799994)\ (5.585452)}} = .544271.$$

The value $r_{12\cdot 3}$ as determined by the classical method is

$$r_{12\cdot 3} = \frac{\Delta_{12}}{\sqrt{\Delta_1 \Delta_2}} = \frac{.5016}{.9216} = .544271.$$

Of course, as in the case of multiple correlation, it is better to use classical methods if as few as three variables are involved. The advantage of the factorial method increases with the number of variables involved and particularly so if many partial correlations are desired. The flexibility, too, should be noted since it is possible to obtain the partial correlation between two variables against any subset of the original variables.

5. *The case of a single factor.* In the case of a single factor,

$$\Delta[I + BB'] = 1 + \sum B^2_{i1}, \quad B_j = B_{j1}, \quad -B'_i = -B_{i1},$$

$$\Delta_i[I + BB'] = 1 + \sum B^2_{k1} - B^2_{i1},$$

$$\Delta_j[I + BB'] = 1 + \sum B^2_{k1} - B^2_{j1}; \text{ so that}$$

$$r_{ij\cdot 12\cdots n} = (-1)^{i+j+1} \sqrt{\frac{B^2_{i1} B^2_{j1}}{(1 + \sum B^2_{k1} - B^2_{i1})(1 + \sum B^2_{k1} - B^2_{j1})}}. \quad (8)$$

From the one-factor problem used in Table III we note that $B^2_{i1} = 1.415$, $B^2_{j1} = 1.203$ and, if we consider the first three variables only, $1 + B^2_{21} + B^2_{31} = 3.252$, $1 + B^2_{11} + B^2_{31} = 3.464$, and

$$r_{12\cdot 3} = \sqrt{\frac{(1.415)(1.203)}{(3.252)(3.464)}} = .389.$$

This is in error because the residuals were not exactly 0. The actual value of $r_{12\cdot 3}$, as found by conventional methods, is .428. The absolute value of the difference is .039 and is about the same size as the average absolute value of the size of the three residuals (5, p. 148), which is .035.

It can be shown in general from (8) that, in the case of one factor, the partial correlation approaches 0 as the number of variables increases. The old factor loadings remain the same with each new addition and if periodically there is added at least one factor loading as large as ε, the denominators increase while the numerators remain the same.

The ease with which partial correlations of high order are computed, in the case of one factor, is shown in Table I where the values of r_{12}, $r_{12\cdot 3}$, $r_{12\cdot 34}$, $r_{12\cdot 345}$, $r_{12\cdot 3456}$, $r_{12\cdot 34567}$, $r_{12\cdot 345678}$, are exhibited.

TABLE I

Computation of the partial correlations $r_{12\cdot}$.

Variable	a_{j1}^2	u_j^2	B_{j1}^2	$B_{11}^2 B_{21}^2$	$r_{12\cdot}$	$M_{11)1(}$	$M_{11)2(}$	$M_{11)1(}$ / $M_{11)2(}$	$r_{12\cdot}^2$	$r_{12\cdot}$
1	.586	.414	1.415	1.702						
2	.546	.454	1.203	1.702		2.203	2.415	5.320	.320	.566
3	.512	.488	1.049	1.702	.3	3.252	3.464	11.265	.151	.389
4	.451	.549	.821	1.702	.34	4.073	4.285	17.453	.0975	.312
5	.402	.598	.672	1.702	.345	4.745	4.957	23.521	.0724	.269
6	.357	.643	.555	1.702	.3456	5.300	5.512	29.214	.0583	.241
7	.354	.646	.548	1.702	.34567	5.848	6.060	35.439	.0480	.219
8	.331	.669	.495	1.702	.345678	6.343	6.555	41.578	.0409	.202

The 1764 different partial correlation coefficients of this problem can be computed in this way.

6. *Alternate form of the general formula.* The numerator of (7) can be expressed as the sum of two determinants of the rth order rather than one of the $(r+1)$st order. Since

$$\Delta\begin{bmatrix} I & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = \Delta[I+BB'] + \Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix}$$

we have

$$\begin{aligned}\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} &= -\Delta[I+BB'] + \Delta\begin{bmatrix} I & B_2 \\ -B'_1 & I+BB' \end{bmatrix} \\ &= -\Delta[I+BB'] + \Delta\begin{bmatrix} I & B_2 \\ 0 & I+BB'+B_2B'_1 \end{bmatrix} \\ &= -\Delta[I+BB'] + \Delta[I+BB'+B_2B'_1]\,. \end{aligned} \tag{9}$$

For example, in the illustration of section 4,

$\Delta[I+BB'] = 15.094540$, while

$$\Delta[I+BB'+B_2B'_1] = \begin{bmatrix} 7.148929 & .395329 \\ .189358 & 2.765149 \end{bmatrix} = 19.696570\,,$$

and

$$\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = 4.602030\,,$$ as indicated earlier.

7. *Approximation Formulas.* The computation can be considerably simplified if we make an approximation by placing the nondiagonal elements of $[I+BB']$ equal to zero. In this case the matrix $[I+BB']$ is a diagonal matrix with elements M_{kk}. Then

$$\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = \sum_{k=1}^{r} B_{1k}B_{2k}\,\Pi\frac{M_{ii}}{M_{kk}}\,, \quad \text{and}$$

$$r_{12\cdot34\cdots n} = \frac{\sum B_{1k}B_{2k}\,\Pi\, M_{ii}/M_{kk}}{\sqrt{\Pi\, M_{ii)1(}\;\Pi\, M_{ii)2(}}}\,. \tag{10}$$

As an illustration, we compute an approximation to $r_{12\cdot34567}$ for the two-factor problem of section 4. In this case, if we take the factor loadings to three places we get $B_{11}B_{21} = 1.506$, $B_{12}B_{22} = .0885$, $M_{11} = 11.753$, $M_{22} = 5.266$, $M_{11)1(} = 10.963$, $M_{11)2(} = 8.883$, $M_{22)1(} = 5.239$, $M_{22)2(} = 4.971$, with

$$r_{12\cdot34567} = \frac{(1.506)\,(5.266) + (.0885)\,(11.753)}{\sqrt{(10.963)\,(8.883)\,(5.239)\,(4.971)}} = .178\,.$$

A further approximation can be made, if n is large enough so

that M_{ii}, $M_{ii)1(}$ and $M_{ii)2(}$ are approximately equal. In this case

$$r_{12\cdot34\cdots n} = \sum_{k=1}^{r} \frac{B_{1k}\,B_{2k}}{M_{kk}}. \tag{11}$$

It can be shown that $r_{12\cdot34\cdots n} \to 0$ as $n \to \infty$ if periodically some $a_{ij} > \varepsilon$, since B_{1k} and B_{2k} remain the same, but M_{kk} increases. See also (12, section 15).

8. *Regression coefficients.* From (2) and (5) it is apparent that

$$\beta_{12\cdot34\cdots n} = \frac{\Delta_{12}}{\Delta_1} = \frac{u_1u_2u_3^2\cdots n_n^2\,\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix}}{u_2^2u_3^2\cdots u_n^2\,\Delta_1[I+BB']}$$

$$= \frac{u_1\,\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix}}{u_2\qquad \Delta_1[I+BB']}.$$

In general, similarly

$$\beta_{ij\cdot12\cdots n} = (-1)^{i+j+1}\frac{u_i\,\Delta\begin{bmatrix} 0 & B_j \\ -B'_1 & I+BB' \end{bmatrix}}{u_j\,\Delta_i[I+BB']}. \tag{12}$$

Approximations can be made as in the last section. See also (12, section 7).

PART III

The Different Multiple Correlations Involving the Observed Variables and the Common Variables

1. *Introduction.* In this part a study is made of the multiple correlation coefficient (1) when an observed variable is estimated from (a) common variables, (b) common variables and observed variables, (c) observed variables; (2) when a common variable is estimated from (a) common variables, (b) common variables and observed variables, and (c) observed variables.

In this part, we make use of the fact that the correlation between an observed variable x_j and a common variable y_i is equal to a_{ji}. This well-known fact follows at once from [Part I, Equation (1)] by forming

$$\frac{\sum x_j\,y_i}{N} = \frac{\sum a_{ji}\,y_i^2}{N} = a_{ji}. \tag{1}$$

The matrix of the correlations of the n observed variables and the r common variables can be written in the form $\begin{bmatrix} R & A \\ A' & I \end{bmatrix}$ if the observed variables precede the common variables horizontally to the right and vertically downward.

2. *The multiple correlations of an observed variable.* The multiple correlations of the observed variables have been worked out previously. Results and references are here stated.

(a) *Correlation with the common variables.*

The multiple correlation of the observed variable x_j with the r common variables has been shown by various authors (7, p. 2) (10, p. 164) to be equal to h_j.

(b) *Correlations with the common variables and the* n — 1 *observed variables.*

It is likewise well-known (7, p. 4) (10, p. 165) that the multiple correlation of an observed variable with the r common variables and the n — 1 observed variables is equal to h_j.

c) *Correlation with the* n — 1 *observed variables.*

It is shown in Part I that

$$R_j = \sqrt{1 - u_i^2 \frac{\Delta[I + BB']}{\Delta_j[I + BB']}}. \tag{2}$$

It is not necessary to use all the n — 1 other variables in applying this formula.

3. *The multiple correlations of the common variable.*

(a) *With the* r — 1 *other common variables.*

This trivial case results in $R_i = 0$, since $\Delta = \Delta_i = 1$ in the formula (4, p. 301).

$$R_i = \sqrt{1 - \frac{\Delta}{\Delta_i}}. \tag{3}$$

(b) *With the* r — 1 *other common variables and the* n *observed variables.*

The value of Δ in the determinantal formula (3) is now

$$\Delta \begin{bmatrix} R & A \\ A' & I \end{bmatrix} = \Delta[U^2]$$

as shown in Part I. The value of Δ_i is not quite so easily obtained since it calls for the deletion of row $n + i$ and column $n + i$ in the matrix having $n + r$ rows and $n + r$ columns. Now the order in which the y's occur is not material so, for convenience, we assume that y_i is the last variable y_r. We then write the factorial matrix in

two parts. The first part, which we designate $A_{)r(}$, indicates the matrix composed of all the columns of the factorial matrix with the exception of column r. The second part of the matrix, which we designate by A_r, consists of the rth column only. In this notation the matrix of the intercorrelations of the $n + r$ variables becomes

$$\begin{bmatrix} R & A_{)r(} & A_r \\ A'_{)r(} & I & 0 \\ A'_r & 0 & I \end{bmatrix}$$

while the determinant $\underline{\Delta}_r$ can be written

$$\Delta \begin{bmatrix} R & A_{)r(} \\ A'_{)r(} & I \end{bmatrix} = \Delta \begin{bmatrix} R & A_{)r(} & A_r \\ A'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix} = \Delta \begin{bmatrix} U^2 & A_{)r(} & A_r \\ 0 & I & 0 \\ -A'_r & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\, \Delta \begin{bmatrix} I & B_{)r(} & B_r \\ 0 & I & 0 \\ -B'_r & 0 & I \end{bmatrix} = \Delta[U^2]\, \Delta \begin{bmatrix} I + B_r B'_r & B_{)r(} & B_r \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\,(1 + B^2_{1r} + B^2_{2r} + \cdots + B^2_{nr})$$

$$= \Delta[U^2]\,\left(1 + \frac{a^2_{1r}}{u^2_1} + \frac{a^2_{2r}}{u^2_2} + \cdots + \frac{a^2_{nr}}{u^2_n}\right). \qquad (4)$$

It follows at once, since i may be substituted for r, that

$$R_i = \sqrt{1 - \frac{\Delta}{\Delta_i}} = \sqrt{1 - \frac{1}{1 + \sum B^2_{ji}}} = \sqrt{\frac{\sum B^2_{ji}}{1 + \sum B^2_{ji}}}. \qquad (5)$$

This formula is also given by Guttman (12, 9a). It is clear that $R_i \leq 1$ and that $R_i \to 1$ as $n \to \infty$ if periodically some B^2_{ji} is at least of size ε.

The formula (5) is adapted to easy computation. It is to be noted that the quantities B_{ji} are quantities which are used in finding the multiple correlation of an observed variable with the other observed variables.

(c) *With* n *observed variables.*

We assume again, without loss of generality, that the common variable is the last common variable, r. We then have

$$R_r = \sqrt{1 - \frac{\Delta}{\Delta_r}} \quad \text{with } \Delta = \Delta \begin{bmatrix} R & A_r \\ -A'_n & I \end{bmatrix} \text{ and}$$

$\Delta_r = \Delta[R] = \Delta[U^2]\Delta[I + BB']$ as shown in Part I. Now

$$\Delta \begin{bmatrix} R & A_r \\ -A'_r & I \end{bmatrix} = \Delta \begin{bmatrix} R & A_{)r(} & A_r \\ 0 & I & 0 \\ -A'_r & 0 & I \end{bmatrix} = \Delta \begin{bmatrix} U^2 & A_{)r(} & A_r \\ -A'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\Delta\begin{bmatrix} I & B_{)r(} & B_r \\ -B'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix} = \Delta[U^2]\Delta[I + B_{)r(}B'_{)r(}] .$$

It follows that, in general,

$$R_i = \sqrt{1 - \frac{\Delta[I + B_{)i(}B'_{)i(}]}{\Delta[I + BB']}} . \tag{6}$$

The reader may refer to Guttman (12, Section 10).

Here $\Delta[I + B_{)i(}B'_{)i(}]$ differs from $\Delta[I + BB']$ only in having row i and column i deleted. As an approximation formula we may set all non-diagonal elements in row i and column i of $I + BB' = 0$ and we get

$$R_i = \sqrt{1 - \frac{1}{1 + \sum B^2_{ji}}} \tag{7}$$

which agrees with the result (5) of the previous case.

4. *Illustration.* As an illustration we take a problem used by Thurstone (5, p. 128) and compute the values of the B's to three significant places. These are presented in Table I.

TABLE I

The Computation of the B's.

(1) Var. j	(2) a_{j1}	(3) a_{j2}	(4) h_j^2	(5) u_j^2	(6) u_j	(7) B_{j1}	(8) B_{j2}
1	.660	.121	.450	.550	.742	.889	.163
2	.830	.266	.760	.240	.490	1.694	.543
3	—.541	.638	.700	.300	.548	—.987	1.164
4	—.126	.771	.610	.390	.625	—.202	1.234
5	.437	.591	.540	.460	.678	.645	.872
6	.638	—.337	.520	.480	.693	.921	—.486
7	.904	.109	.830	.170	.412	2.194	.265

It follows at once that

$$1 + \sum_{j=1}^{7} B^2_{j1} = 11.753 ,$$

$$\sum_{j=1}^{7} B_{j1} B_{j2} = .363 ,$$

$$1 + \sum_{j=1}^{7} B^2_{j2} = 5.266 ,$$

so that, if y_1 is correlated with y_2 and the seven observed variables we get

$$R_{y_1} = \sqrt{\frac{10.753}{11.753}} = .956 .$$

In a similar fashion

$$R_{y_2} = \sqrt{\frac{4.266}{5.266}} = .900 .$$

If y_1 is correlated with the seven observed variables we have

$$R_{y_1} = \sqrt{1 - \frac{5.266}{\begin{vmatrix} 11.753 & 3.63 \\ .363 & 5.266 \end{vmatrix}}} = .956$$

and, similarly

$$R_{y_2} = \sqrt{1 - \frac{11.753}{\begin{vmatrix} 11.753 & 3.63 \\ .363 & 5.266 \end{vmatrix}}} = .900 .$$

The correlation of each observed variable with the remaining observed variable has been exhibited in Part I. A general summary of results is given in Table II.

TABLE II

Different Multiple Correlations.

Predicted Variable	Predicting Variables (Excluding predicted Variable)		
	Common Variables	Common Variables and Observed Vars.	Observed Variables
x_1	.671	.671	.639
x_2	.872	.872	.814
x_3	.837	.837	.731
x_4	.781	.781	.668
x_5	.735	.735	.666
x_6	.721	.721	.673
x_7	.911	.911	.841
y_1	0	.956	.956
y_2	0	.900	.900

6. *Additional Cases.* Cases may arise in which one does not desire to use all the observed variables nor all the common variables. It is possible, in general, to work these cases by means of matrix developments similar to those of this paper. The desired matrix is set up and then built up so that complete correlation and factorial matrices result. The determinant of this matrix is then reduced to the determinant of a matrix having but r rows and r columns.

Frisch (**11**, p. 15) has used "hollow" determinants in which the entries in the diagonals are replaced by 0. These also can be evaluated

from the factorial matrix, but the loadings are now "communality augmented" rather than "uniqueness augmented" loadings. For

$$\Delta[R-I] = \Delta\begin{bmatrix} R-I & A \\ 0 & I \end{bmatrix} = \Delta\begin{bmatrix} U^2-I & A \\ -A' & I \end{bmatrix} = \Delta\begin{bmatrix} -H^2 & A \\ -A' & I \end{bmatrix}$$
$$= (-1)^n \Delta[H^2]\, \Delta\begin{bmatrix} I & C \\ C' & I \end{bmatrix} = (-1)^n \Delta[H^2]\, \Delta[I-CC']\,. \tag{8}$$

where $\Delta[H^2] = h^2_1 h^2_2 \cdots h^2_n$ and the elements of C are

$$C_{ji} = \frac{a_{ji}}{h_j}\,. \tag{9}$$

7. *Conclusion.* It has been shown in this paper how the results of a factorial matrix can be used in obtaining multiple and partial correlations. The presentation is illustrative rather than exhaustive, general rather than rigorous, but it does emphasize the actual computation of multiple and partial correlations in problems involving many variables. Special attention is given to formulas which, with a relatively small amount of work, will yield good approximations to the desired quantities.

From the general theoretical point of view, it seems, there is one point which should be emphasized and that is that the correlation matrix results from an incomplete preliminary reduction of the data since the correlation matrix does not permit the isolation of a set of numbers which accompany one variable only. To be sure the correlation matrix does provide a set of numbers for each variable, but these numbers are associated with other variables; in fact, there has to be one of these numbers for each of the other variables. However, if the correlation matrix is reduced to a factorial matrix, we have each variable indicated by a set of numbers, numbers which are, in form at least, independent of the other variables, and from these numbers we can now construct not only the correlation matrix, but also the classical partial and multiple correlation coefficients as well. Since the numbers associated with the observed variable are not associated with the other observed variables, the method makes feasible the immediate study of any subset of the observed variables. This flexibility is most important as is the fact that the representation of a variable by such a set of numbers gives promise of the development of very objective criteria for the inclusion or exclusion of a variable in prediction (**10**, **11**, **12**). It appears then that the computation of a factorial matrix is indicated as a proper step in preliminary reduction of

the data by the student of multiple correlation as well as by the student of multiple factor analysis.

REFERENCES

1. Kurtz, A. K. The use of the Doolittle method in obtaining related multiple correlation coefficients. *Pschometrika*, 1936, 1, 45-51.
2. Dwyer, P. S. The simultaneous computation of groups of regression equations and associated multiple correlation coefficients. *Annals Math. Stat.*, 1937, 8, 224-231.
3. Tolley, H. R. and Ezekial, Mordecai. The Doolittle method for solving multiple correlation equations versus the Kelley-Salisbury iteration method. *J. Am. Stat. Ass.*, 1927, 22, 497-500.
4. Kelley, T. L. Statistical method. New York: Macmillan, 1923.
5. Thurstone, L. L. The vectors of mind. Chicago: Univ. Chicago Press, 1935.
6. Wilson, E. B. and Worcester, Jane. Note on factor analysis. *Psychometrika*, 1939, 4, 133-148.
7. Roff, M. Some properties of the communality in multiple factor theory. *Psychometrika*, 1935, 1, 1-6.
8. Dwyer, P. S. The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika*, 1937, 2, 173-178.
9. Dwyer, P. S. An analysis of 19 occupational scores, etc. *J. Appl. Psychol*, 1938, 22, 8-16.
10. Dwyer, P. S. The contribution of an orthogonal multiple factor solution to multiple correlation. *Psychometrika*, 1939, 4, 163-171.
11. Frisch, R. Statistical confluence analysis by means of complete regression systems. Oslo, 1934.
12. Guttman, Louis. Multiple rectilinear prediction and the resolution into components. *Psychometrika*, 1940, 5, 75-99.