PSYCHOLOGICAL RESEARCH

FINAL REPORT

by

George E. Nicholson, Jr., Thomas E. Jeffrey, William G. Howe
Institute of Statistics, University of North Carolina
Chapel Hill, North Carolina

Mimeograph Series No. 100
April 15, 1954
Limited Distribution

TABLE OF CONTENTS

Introduction

The material in this report arose from considering certain methodological problems which arise in the field of test construction and personnel selection. During the past few years various members of the Institute of Statistics research program in statistics have done work on certain aspects of some of these problems and the existence of this work, coupled with the establishment of the Psychometric Laboratory at Chapel Hill in 1952, seemed to make a survey of what had been done and an attempt to relate this with problems of interest to research workers in the field of mental testing desirable.

It was the original intention of this project to formulate outstanding problems in the field of statistical problems of mental testing, survey the theory for what had been accomplished and relate this theory to the problems by means of expository extension and example. In this way it was hoped that statisticians might be made aware of problems in a manner which would command their attention, workers in the testing field would be made familiar with new methods and computational techniques would be developed for the utilization of techniques well known but not extensively applied because of laborious figure work.

This report is principally the work of Professors G. E. Nicholson, Jr., T. E. Jeffrey and Mr. William G. Howe. Some of the topics were discussed with Drs. Dan Teichroew and Masil B. Danford. Dr. Ledyard Tucker of the Educational Testing Service was generous in discussing several points and, together with other members of the Educational Testing Service, supplied data for examples. Dr. Hubert Brogden and Mr. Harry Harman of the Adjutant General's Office also discussed the progress of the report and supplied data from the Personnel Research Branch files. Finally, considerable reference was made to unpublished thesis results of G. E. Nicholson, Jr., D. Teichroew and M. B. Danford.

1. Prediction in Future Samples

A problem which arises in the field of psychology as well as other fields is the following.

Suppose a test battery consisting of p aptitude tests (predictors) is applied to a group of students who are being considered for admission to a particular school. The battery is intended to furnish the admissions officer with information to assist him in predicting the success of the applicants in their first year. After the students have completed the first year a criterion measure (average academic grade, say) is obtained for each student. The test battery is then validated, i.e., the multiple correlation coefficient between test score and criterion is computed. A prediction equation is obtained by using the regression coefficients calculated on a least squares basis as weights to be applied to the test scores of succeeding students. This gives rise to an estimate of the future criterion score of the student and in this sense is useful for predicting the probable success or failure of members of succeeding classes.

When the criterion scores of the second group of students are available, these may be compared with the estimates in order to determine how well the predicting equation is performing. The determination of the "goodness" of the predicting equation is usually made on the basis of the multiple correlation coefficient obtained in the first (validation) sample. However, the object of the procedure is to predict the criterion scores in the second sample and the correlation between the predicted scores and the observed scores in the second sample and the correlation between the predicted scores and the observed scores in the second sample can be regarded as the actual measure of the "goodness" of the prediction equation. It has been observed by several writers that generally the second multiple correlation coefficient is smaller than the first and this effect has been called the "shrinkage" of the multiple correlation coefficient.

To fix ideas and establish a notation, let us consider a set of variates $y, x_1, \ldots, x_p$ where y is the dependent or criterion variate and $x_1, \ldots, x_p$ are the independent or predictor variates. Let $y_\beta(1)$ denote the β-th observation on y and $x_{i\beta}$ the β-th observation on $x_i (i=1,2,\ldots,p;\ \beta=1,2,\ldots,N_1)$ in a sample S_1. Let $y_\alpha(2)$, $x_{i\alpha}(2)$ $(i=1,2,\ldots,p_j, \alpha=1,2,\ldots,N_2)$ be defined for an independent sample S_2. We assume that x_i are fixed constants. In particular let $x_{1\beta}(1)=x_{1\alpha}(2)=1$ and $x_{j\beta}(1)$, $x_{j\alpha}(2)$ $(j=2,\ldots,p)$ be deviations from sample means in S(1) and S(2) respectively. The observations y are random variables such that

(1) $y_1(1), \ldots, y_{N_1}(1), y_1(2), \ldots, y_{N_2}(2)$ are normally and independently distributed with a common unknown variance σ^2.

(2) The expected value of $y = \beta_1 x_1 + \ldots + \beta_p x_p$ where β_i are unknown constants. This implies that $y_\gamma = \eta_\gamma + \varepsilon_\gamma = \beta_1 x_{1\gamma} + \beta_2 x_{2\gamma} + \ldots + \beta_p x_{p\gamma} + \varepsilon_\gamma$ where ε_γ is a normal variate zero mean, and such that $E(\varepsilon_i\ \varepsilon_j) = \sigma^2$ if i=j and zero otherwise.

$$\text{Let } A_{ij} = \underset{\gamma}{S}\, x_{i\gamma} x_{j\gamma};\ A = [a_{ij}];\ C = [c_{ij}] = A^{-1}$$

and

$$Y_\alpha = b_1 x_{1\alpha} + \ldots + b_p x_{p\alpha}$$

where the b_i are defined by the p normal equations

$$s_\alpha (y_\alpha - Y_\alpha)\, x_{i\alpha} = 0$$

Denote by A(1) and A(2) respectively the two matrices of normal equations which define the two sets of regression coefficients with which we are concerned, i.e.,

$$A(1) = [a_{ij}(1)] = [Sx_{i\alpha}(1)\, x_{j\alpha}(1)]$$

and similarly for A(2). Also $A(1)/N_1$ is the sample covariance matrix for the little x's in sample one and a similar interpretation holds for $A(2)/N_2$. Now consider the following 2 x 2 table.

	Sample 1	Sample 2
Y(1)	R^2_{11}	R^2_{12}
Y(2)	R^2_{21}	R^2_{22}

In the above table Y(1) represents the least squares predicting equation calculated from sample 1 and Y(2) represents the least squares predicting equation derived from sample 2. R^2_{12} is the result of correlating the estimated criterion variates in S_2 on the basis of the prediction equation derived from S_1.

It is clear that in this notation R^2_{11} is simply the square of the multiple correlation coefficient claculated from the 1-st sample and R^2_{12} is the square of the correlation between the observed criterion values in the 2-nd sample and those estimated from the predictors in the 2-nd sample and the regression coefficients calculated from the 1-st sample. The problem which we consider arises from the observation of research workers that generally R^2_{12} is lower than R^2_{11}.

If sample 1 and sample 2 are independently drawn from the same population, then R^2_{11} and R^2_{22} are independent estimates of the population parameter ρ^2. Under the assumption already made the distribution of these sample estimates is well known. Accordingly questions related to the behaviour of the best linear prediction to be expected in any sample may be answered by using this distribution.

For example, the chance that R^2_{22} will be less than R^2_{11} is 1/2 since both follow the same distribution. The chance that $R^2_{11} - R^2_{22} > d$ may be answered by working directly with the known distribution.

The distribution of R^2_{ii} depends on N_i, the number of observations in the sample, p, the number of predictors, and ρ^2, the parameter. Previous considerations of this problem have resulted in proposals to correct the observed R^2_{ii} for bias to obtain a better estimate of ρ^2. A first order approximation to an unbiased estimate of ρ^2 is

$$\rho^2 = 1 - \frac{N_i - 1}{N_i - p} [1 - R^2_{ii}] .$$

More exact expressions are available for obtaining unbiased estimates of ρ^2 but the problem of inferring how well a predicting equation will perform in a new sample is not answered by this kind of analysis.

The set of regression coefficients estimated from the first sample are known to be unbiased estimates of the population regression coefficients. For example, let $\beta = [\beta_1, \beta_2, \ldots, \beta_p]$ be the vector of population regression coefficients and let $b_1 = [b_{11}, b_{21}, \ldots b_{p1}]$ be the vector of regression coefficients estimated on the basis of sample 1. Then $E(b_1) = \beta$. If the first sample is large then it may be expected that $\max_i (b_i - \beta_i) = \delta$ is very small and it may be considered that this problem reduces itself to considering the sampling fluctuations of the simple correlation coefficient since for a fixed set of β the multiple correlation coefficient simply reduces to the correlation between the criterion and a fixed linear function of normally distributed variables which is a single normal variable. Then the problem becomes that of studying the distribution of the ordinary correlation coefficient. The question, however, remains of why there should be any shrinkage since it would be expected that greater values of the second coefficient would occur as frequently as smaller ones. The intuitive answer seems to run as follows. If the first sample is large and has ample degrees of freedom for estimating β then it is reasonable to assume that the regression weights are close to the population weights. Since b is the unique set of coefficients which maximizes the multiple correlation coefficient in the sample, the observed R_{11}^2 is an over estimate of ρ^2. In a second sample, however, of approximately the same size the b used being close to the β of the population will tend to make R_{12}^2 an unbiased estimate of ρ^2 and hence less than R_{11}^2 which is biased upward. Accordingly, the phenomenon of shrinkage should be observed whenever two fairly large samples are used and when the degrees of freedom available for estimation of β are large. Furthermore, in this case the phenomenon ought to be adequately explained by appealing to the distribution of the simple correlation coefficient. The problem, however, will depend on the number of degrees of freedom available for estimating β in sample 1 and the size of the second sample relative to the first.

Let us consider the following formulation of the problem. The best prediction in a particular sample is obtained when the least squares prediction equation derived from that sample is used. A measure of the adequacy of this prediction equation is given by R_{11}^2. The actual measure of prediction efficiency is, however, R_{12}^2 since it is the second sample in which the prediction is required and not the first. Now among the various interpretations which can be made of the correlation coefficient, one is that the square of the correlation coefficient is the proportion of variance of one of the variables which is explained by a linear least squares regression on the other. It is this interpretation which we adopt. We propose that the accuracy of a predicting equation in a sample be measured by this means. If then the y_α in a sample of N_i

observations are to be estimated by an estimation procedure $Y(\alpha)$ which gives rise to N_i estimates $Y_\alpha(k)$ then we shall define the correlation between y and $Y(k)$ to be

$$R^2(I) = 1 - \frac{\underset{\alpha}{S}(y_\alpha - Y_\alpha(k))^2}{\underset{\alpha}{S}(y_\alpha - \bar{y})^2}$$

It should be noted that the relation

$$\frac{[\underset{\alpha}{S}(y_\alpha - \bar{y})(Y_\alpha(k) - \bar{Y}(k))]^2}{\underset{\alpha}{S}(y_\alpha - \bar{y})^2 \cdot \underset{\alpha}{S}(Y_\alpha(k) - \bar{Y}(k))^2} = 1 - \frac{\underset{\alpha}{S}(y_\alpha - Y_\alpha(k))^2}{\underset{\alpha}{S}(y_\alpha - \bar{y})^2}$$

is not true unless $Y_\alpha(k)$ is the estimate based on the least squares regression calculated from the same sample. We, therefore, select as our definition

$$(1) \qquad R_{12}^2 = 1 - \frac{\underset{\alpha}{S}(y_\alpha(2) - Y_\alpha(1))^2}{\underset{\alpha}{S}(y_\alpha(2) - \bar{y}(2))^2}$$

The failure of $Y_\alpha(1)$ to be equal to $y_\alpha(2)$ is measured by

$$S(y_\alpha(2) - Y_\alpha(1))^2 = S(y_\alpha(2) - Y_\alpha(2))^2 + S(Y_\alpha(2) - Y_\alpha(1))^2$$

which may be partitioned into the two components shown. The first term represents the failure of the observations to follow the best least squares predicting equation and the second part represents the failure of the predicting equation derived from Sample 1 to be the same as that calculated from Sample 2. It, therefore, seems natural to use the ratio of $S[y_\alpha(2) - Y_\alpha(1)]^2$ to $S[y_\alpha(2) - Y_\alpha(2)]^2$ as a measure of the efficiency of the predictor $Y_\alpha(1)$ in Sample 2. Let

$$E = \frac{\underset{\alpha}{S}[y_\alpha(2) - Y_\alpha(2)]^2}{\underset{\alpha}{S}[y_\alpha(2) - Y_\alpha(1)]^2}$$

measure the efficiency of the predicting equation. It is clear that $E \leq 1$ and assumes its maximum value for $Y_\alpha(1) = Y_\alpha(2)$. Also from (1)

$$E = \frac{(1 - R^2_{12})\underset{\alpha}{S}(y_\alpha(2) - \bar{y}(2))^2 - S(Y_\alpha(2) - Y_\alpha(1))^2}{S(y_\alpha(2) - Y_\alpha(1))^2}$$

If $Y_\alpha(1)$ is taken as $\bar{y}(2)$ we see that $E < 1 - R^2$ showing that for a large R^2 using a simple predicting equation like the mean of the y's has very low efficiency.

It may be emphasized here that in terms of our assumptions the problem of finding a set of constants $b_1, \ldots, b_p$ which will maximize

$$\frac{E[y - E(y)][\Sigma_1^p b_i x_i - E(\Sigma_1^p b_i x_i)]^2}{E[y - E(y)]^2 \, E[\Sigma_1^p b_i x_i - (E\Sigma_1^p b_i x_i)]^2} = r^2$$

is the same as the problem of finding constants $a_1, \ldots, a_p$ such that

$$E(y - \Sigma_1^p a_i x_i)^2$$

is a minimum, i.e., $a_i = b_i$ and

$$1 - \frac{E(y - \Sigma_1^p b_i x_i)^2}{E[y - (Ey)]^2} = r^2 \quad .$$

However this relationship is not true for an arbitrarily determined set of constants. The definition of multiple correlation coefficient must then be chosen, made specific and we have chosen what is sometimes referred to as the correlation index.

$$S[y_\alpha(2) - Y_\alpha(1)]^2 = S(y_\alpha(2) - Y_\alpha(2))^2 + \sum_{ji}\sum [b_i(2) - b_i(1)]$$

$$[b_j(2) - b_j(1)]a_{ij}(2) = Q_1 + Q_2 \ .$$

It follows from all the assumptions that Q_1 is distributed like $\chi^2_{N_2 - p}$ and is independent of Q_2. If we let $b_i(2) - b_i(1) = w_i$ it follows that the set w_i has a normal multivariate distribution with zero means and covariance matrix $\Sigma = [C(1) + C(2)]$. This covariance matrix Σ and the matrix of the form $Q_2 = A(2)$ are both positive definite. Therefore Q_2 may (by a familiar transformation) be transformed into the form

$$Q_2 = \lambda_1 z_1^2 + \ldots + \lambda_p z_p^2$$

where the real positive λ_i are the roots of the determinantal equation

$$|A(2)\ C(1) - (\lambda - 1)| \ = 0$$

and the z_i are normally distributed random variables with zero means and unit variance. The final result then is

$$\frac{Q_2}{Q_1} = 1 + \frac{\sum_i \lambda_i z_i^2}{\chi^2_{N_2-p}}$$

We now consider the reciprocal of the ratio $\dfrac{S(y_\alpha(2) - Y_\alpha(2))^2}{S(y_\alpha(2) - Y_\alpha(1))^2} = E$

which is

$$\frac{1}{E} = \frac{Q_1+Q_2}{Q_1} = 1 + \frac{Q_2}{Q_1} = 1 + \frac{Q_2}{\chi^2_{N_2-p}}$$

and

$$P = \frac{1}{E} - 1 = \frac{Q_2}{\chi^2_{N_2-p}} \quad .$$

Special Cases.

Assume that $b_i(1) = \beta_i$, i.e., the first sample is so large that the estimates of the β_i are equal to the β_i. In this case Q_2 is distributed exactly as χ^2_p and

$$\Pr\left[\frac{1}{E} \leq x\right] = \Pr\left[1 + \frac{\chi^2_p}{\chi^2_{N_2-p}} \leq x\right] = \Pr\left[F_{p,N_2-p} \leq \frac{N_2-p}{p}(x-1)\right]$$

where F_{p,N_2-p} is Snedecor's F with p and N_2-p degrees of freedom.

Assume that $A(2) = A(1)/k$ and consider the determinantal equation

$$|A(2)C(1)-(\lambda-1)| = 0 .$$

Under this assumption all of the roots are equal, $\lambda_i = k+1$ and $Q_2 = (1+k)\chi_p^2$. Then

$$\Pr\left[\frac{1}{E} \leq x \right] = \Pr\left[F_{p,N_2-p} \leq \frac{(x-1)(N_2-p)}{p(1+k)} \right] .$$

In order to obtain an upper confidence limit for P at the ε level of significance, choose x such that it satisfies the equation

$$F_\varepsilon = \frac{(x-1)(N_2-p)}{p(1+k)}$$

that is

$$x = 1 + \frac{p(1+k)}{N_2-p} F_\varepsilon$$

where

$$\Pr\left[F_{p,N_2-p} \leq F_\varepsilon \right] = 1 - \varepsilon \quad .$$

The confidence interval is

$$1 \leq \frac{1}{E} \leq x$$

with confidence coefficient $1 - \varepsilon$.

Since E is the relative efficiency of the prediction equation Y(1) derived from S(1) the interpretation of the confidence interval gives an estimate of how well Y(1) will predict in S(2). The practical significance of this result is obvious for large samples $A(2)/N_2$ should be approximately equal to $A(1)/N_1$.

For $k = N_2/N_1$

$$\frac{1}{E} \approx 1 + \frac{(1 + \frac{N_2}{N_1})\ \chi^2_p}{\chi^2_{N_2-p}}$$

The expected value of $\frac{1}{E}$ is

$$E(\frac{1}{E}) = E \left[1 + (1 + \frac{N_2}{N_1})\ F_{p,N_2-p} \cdot \frac{p}{N_2-p-2} \right]$$

$$= \frac{N_1(N_2-2) + N_2 p}{N_1(N_2-p-2)}$$

which for large N_2 is approximately

$$E(\frac{1}{E}) = \frac{1+ \frac{p}{N_1}}{1- \frac{p}{N_2}} \quad .$$

We note from this that if p is fixed and N_1 and N_2 are both large then $E(\frac{1}{E}) = 1$ which indicates that the regression equation calculated from a very large sample will predict equally well in another very large sample provided the number of predictors is relatively small. Let $N_2 = rN_1$ and $p = sN_1$ where $0 \leq s \leq 1$ and $r > 0$. Then

$$E(\frac{1}{E}) = \frac{(1+s)r}{(r-s)}$$

and

$$\frac{1}{E(\frac{1}{E})} = \frac{r-s}{r(1+s)} \quad .$$

For this expression to have its maximum value s must be a minimum. Hence we have the conclusion above that the ratio of the number of predictors to the sample size N_1 should be small. However if r is very large for small s we have an efficient arrangement. For $A(2) = \frac{N_2}{N_1} A(1)$ then the following table gives the expected relative efficiency of the predictor Y(1) in S_2.

N_2/N_1 \ p/N_1	.01	.05	.1	.5	1
.1	.89	.47			
.2	.94	.71	.45		
.5	.97	.86	.73		
1	.98	.90	.82	.33	
2	.99	.93	.86	.50	.25
10	.99	.95	.90	.63	.45
oo	.99	.95	.91	.66	.50

Another possible measure of prediction is

$$\frac{1}{E'} = \frac{N_1}{N_2} \frac{S(y_\alpha(2) - Y_\alpha(1))^2}{S(y_\beta(1) - Y_\beta(1))^2} .$$

This is the ratio of the sample variance of the residuals about the predictor from S(1) applied in S(2) to the predictor of S(1) applied in S(1). By the same arguments as before $\frac{1}{E'}$ may be expressed as

$$\frac{N_1}{N_2} \quad \frac{\chi^2_{N_2-p} + \sum_{i=1}^{p} \lambda_i z_i^2}{\chi^2_{N_1-p}}$$

where $\chi^2_{N_1-p}$ is independent of the numerator and the quantities in the numerator have already been defined. If we assume $A(1) = \frac{A(2)}{k}$ by the previous argument we have

$$\frac{1}{E'} = \frac{N_1}{N_2} \; \frac{\chi^2_{N_2-p} + (1+k)\,\chi^2_p}{\chi^2_{N_1-p}}$$

and therefore

$$E\left(\frac{1}{E'}\right) = \frac{N_1}{N_2} \cdot \frac{N_2+kp}{N_1-p-2}$$

from which it follows that if $k = \frac{N_2}{N_1}$

$$E\left(\frac{1}{E'}\right) = \frac{1 + \frac{p}{N_1}}{1 - \frac{p+2}{N_1}}$$

It is of interest to note this expression is independent of N_2.

In order to present an empirical example based on the analysis described above two sets of data were used. The first set of data was obtained through the courtesy of Dr. Ledyard Tucker and Marjorie Olsen of the Educational Testing Service, Princeton, N.J., and is from Law

Schools admissions tests as described below. The notation

$$R'_{12} = \frac{S(y_\alpha(1)-\bar{y}(1))(Y_\alpha(2)-\bar{Y}(2))}{\sqrt{S(y_\alpha(1)-\bar{y}(1))^2 S(Y_\alpha(2)-\bar{Y}(2))^2}} \quad \text{and} \quad R^2_{12} = 1 - \frac{S(y_\beta(2)-Y_\beta(1))^2}{S(y_\beta(2)-\bar{y}(2))^2}$$

is used.

Law Schools Admissions Test: Records for 1,000 freshman students.

	School	Sample I	Sample II	Total
1.	Buffalo	44	44	88
2.	Columbia	64	64	128
3.	Cornell	47	47	94
4.	NYU	35	35	70
5.	Penn.	97	97	194
6.	Rutgers	55	55	110
7.	So. Calif.	62	62	124
8.	Virginia	27	27	54
9.	Yale	69	69	138
	Total	500	500	1000

Raw scores for six of the ten subtests of the Law School Admission Test and the total scores over all ten subtests were available. The criterion was the first year average grades converted to a common scale for all schools.

Since Penn. had the greatest number of students, this sample was used. Sample I was used for determining the multiple correlation. The regression weights were determined and used in Sample II, to predict the criterion. Sample II was divided into two parts. The first fifty constituted the first group for which the criterion was predicted and then the entire 97 were used and the criterion predicted.

This was done for three predictors and then for six predictors. The subtests used as predictors are identified as follows:

Part 2 - Sentence Completion
Part 3 - Paragraphs
Part 5 - Reading Comprehension
Part 8 - Figure Classification
Part 7 - Debates
Part 10- Reading Comprehension

Parts 2, 3, and 10 were used as the three predictors while all six were used in the second case.

Results:

N_1	N_2	p	R^2_{11}	R'^2_{12}	R^2_{12}	$\frac{1}{E*} = \frac{S(y-Y(1))^2}{S(y-Y(2))^2}$	$E(\frac{1}{E*})$	Confidence Interval	o/o
97	50	3	.2436	.1256	.1139	1.17	1.0967	1-1.27	95
								1-1.40	99
97	97	3	.2436	.1229	.0697	1.23	1.0638	1-1.17	95
								1-1.26	99
97	50	6	.3345	.2004	.1939	1.21	1.13	1-1.46	95
								1-1.65	99
97	97	6	.3345	.1652	.0771	1.38	1.13	1-1.29	95
								1-1.39	99

In the above calculations the estimate of P is taken to be

$$\frac{1}{E*} = \frac{1 - R^2_{12}}{1 - R^2_{11}}$$

instead of

$$\frac{1}{E} = \frac{1 - R^2_{12}}{1 - R^2_{22}}$$

The following tables give the calculations upon which this example is based.

Law School Admission Test Data

Penn. Two samples of 97 each

Table of Intercorrelations for Sample I

Part 2	-----					
Part 3	.7387	-----				
Part 5	.7162	.6484	-----			
Part 6	.2759	.3995	.3286	-----		
Part 7	.5495	.5192	.5542	.2772	-----	
Part 10	.8230	.7570	.7329	.3321	.5932	-----
Criterion	.4576	.4319	.4651	.4017	.4357	.4672

Raw Score Regression Weights:

(1) For three predictors (Part scores 2, 3, and 10)

b_2 = .1097

b_3 = .0961

b_{10} = .1126

b_0 = 6.4245

Multiple R_{11} = .4919

Predicting criterion in Sample II

N = 50; R'_{12} = .3544

N = 97; R'_{12} = .2497

(2) For six predictors (Part scores 2, 3, 5, 6, 7 and 10)

b_2 = .0878

b_3 = .0005

b_5 = .1055

b_6 = .1642

b_7 = .1204

b_{10} = .0320

b_0 = .9254

Multiple R_{11} = .5784

Predicting criterion in Sample II

N = 50: R'_{12} = .4477

N = 97: R'_{12} = .4064

A second set of data was used in order to illustrate the theory developed. These data were supplied by the Adjutant General's Office and consisted of test scores of a group of 651 EM who attended the 3rd Armored Division Clerks Training School Course at Ft. Knox, Kentucky. These data were divided into six sub groups, each made up of complete classes in terms of starting data. These six sub-groups were:

Sub-Group No.	No. of E. M.
1	113
2	130
3	104
4	97
5	101
6	106

Sub-group 3 (104 cases) was used as one population for determining the regression weights. Sub-groups 3 and 4 (201 cases) and sub-groups 3, 4, and 5 (302 cases) were used as the other two. For each of these three populations the regression weights were obtained using 3, 6, and 8 predictors.

The Predictors:

The predictors are 7 of the 10 aptitude scores derived from the army classification battery. Three of these 10 scores (Army Radio Code Aptitude Score, Electrical Information Score, and Radio Information Score) could not be used since there were so many unsatisfactory or missing scores. Civilian Education was used as a predictor and at first it was intended to use the Adjutant General's Office Aptitude Area IV Score. This score is the sum of three of the other seven scores used. It was later decided against using the Area IV Score and only eight predictors were used in all.

Predictor

No.	Name
1	Civilian Education
2	Reading and Vocabulary Test - RV
3	Arithmetic Reasoning Test - AR
4	Pattern Analysis Test - PA
5	Mechanical Aptitude Test - MA
6	Army Clerical Speed Test - ACS
7	Shop Mechanics Test - SM
8	Automotive Information Test - AI

The Effect on the Multiple Correlation Due to Increasing the Number of Predictors for Three Populations of Different Size

Population No.	Number of Predictors: Three (Nos. 1, 4, 5)	Six (Nos. 1, 3-7 incl.)	Eight (Nos. 1-8 incl.)
I (N = 104)	.5047	.5679	.5897
II (N = 201)	.4997	.5861	.6141
III (N = 302)	.5246	.6378	.6512

Predicting the Criterion:

Since three populations were used to determine the regression weights for each of three groups of predictors, there were nine prediction equations. These are given on the attached work sheets. For each predictor group the criterion was predicted for each of the populations shown below:

Population size used in determining multiple regression weights	Size of population for which criterion is predicted: A	B	C
I (N = 104)	53	106	214
II (N = 201)	101	219	450
III (N = 302)	130	219	349

Regression weights were determined for three different sized groups.

Group I: Sample 3 (104 men)

Group II: Sample 3 plus sample 4 (201 men)

Group III: Sample 3 plus sample 4 plus sample 5 (302 men)

Tables of Intercorrelations for the three groups:

Group I (N = 104)

	1	2	3	4	5	6	7	8
1	---							
2	485	---						
3	454	388	---					
4	156	091	441	---				
5	232	205	378	542	---			
6	150	167	314	309	170	---		
7	166	167	443	434	591	121	---	
8	-065	025	178	297	609	000	548	---
Criterion	470	413	464	229	267	268	255	034

Group II (N = 201)

	1	2	3	4	5	6	7	8
1	---							
2	403	---						
3	353	430	---					
4	099	240	477	---				
5	210	337	476	551	---			
6	165	237	407	315	227	---		
7	126	298	487	449	633	180	---	
8	-007	171	362	365	636	056	587	---
Criterion	423	459	505	283	307	284	333	170

Group III (N = 302)

	1	2	3	4	5	6	7	8
1	---							
2	447	---						
3	359	449	---					
4	128	237	508	---				
5	248	357	485	556	---			
6	224	276	415	375	267	---		
7	166	313	494	450	623	178	---	
8	048	186	401	379	639	101	611	---
Criterion	422	451	583	348	338	341	347	212

Results

Group	N_1	N_2	p	R^2_{11}	R'^2_{12}	R^2_{12}
	104	53	3	.2547	.2515	.2314
	104	106	3	.2547	.2580	.1859
	104	214	3	.2547	.2140	.2134
	104	53	6	.3225	.3841	.3319
1	104	106	6	.3225	.4014	.3566
	104	214	6	.3225	.3257	.3189
	104	53	8	.3477	.3463	.3004
	104	106	8	.3477	.4385	.3975
	104	214	8	.3477	.3641	.3535
	201	101	3	.2497	.2974	
	201	219	3	.2497	.2128	
	201	450	3	.2497	.2485	
	201	101	6	.3435	.5074	
2	201	219	6	.3435	.3078	
	201	450	6	.3435	.3721	
	201	101	8	.3771	.4703	
	201	219	8	.3771	.3626	
	201	450	8	.3771	.3969	
	302	130	3	.2752	.2438	
	302	219	3	.2752	.2373	
	302	349	3	.2752	.2413	
	302	130	6	.4068	.3738	
	302	219	6	.4068	.3437	
	302	349	6	.4068	.3516	
	302	130	8	.4241	.3839	
	302	219	8	.4241	.3827	
	302	349	8	.4241	.3798	

These data give rise to certain unexpected results. In particular it will be noted that in 10 out of the 27 cases computed R'^2_{12} is greater than R^2_{11}. It has already been pointed out that R^2_{22} will be greater than R^2_{12} and that in general R^2_{11} is greater than R^2_{12}. These data were a source of difficulty to analyze and numerous errors and omissions dictated rather restrictively the samples used. As a consequence it is not at all clear that the samples are random samples from the same population and an extensive analysis would have to be undertaken to locate the cause of the unusual behavior of these samples. It was felt that the expenditure of time would not be warranted since the investigators had no direct connection with the original data and as a consequence are in no position to perform such an analysis adequately.

2. The Addition of Tests to a Known Factor Structure

The specific problem deals with relating the Army Classification Battery, consisting of ten tests, to the factor structure obtained by Dr. Adkins in the "Factor Analysis of Reasoning Tests." Personnel Research Section Report No. 878 gives the analysis for this battery of sixty-six tests. These sixty-six tests plus eleven additional tests not included in the analysis, were administered to 200 subjects. The eleventh measure was a criterion score. This was added to the ten tests of the Army Classification Battery. The seventy-seven tests are listed by name in Table 1.

Dwyer's* method was used to obtain the projections of the added tests on the orthogonal factor axes of the reasoning battery. This method depends on the fundamental factor theorem that the orthogonal factor matrix post multiplied by its transpose will reproduce the correlation matrix. Written as an equation, $FF' = R$.

The development of Dwyer's method: consider r tests whose table of intercorrelations, when factored, results in an orthogonal factor matrix of k dimensions. The problem: to find the projections of test t, an additional test, for this factor matrix when the intercorrelations of test t with the other r tests are known. The projections of test t on the k dimensions are defined as: $a_{t1}, a_{t2}, \ldots, a_{tk}$. The following equations must be satisfied since $FF' = R$.

---

*Dwyer, Paul S. "The Determination of the Factor Loadings of a Given Test from the Known Factor Loadings of Other Tests." Psychometrika, Vol. 2, No. 3, September 1937, pp. 173-178.

$$a_{t1}a_{11} + a_{t2}a_{12} + \dots + a_{ti}a_{1i} + \dots + a_{tk}a_{1k} = r_{1t}$$

$$a_{t1}a_{21} + a_{t2}a_{22} + \dots + a_{ti}a_{2i} + \dots + a_{tk}a_{2k} = r_{2t}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_{t1}a_{j1} + a_{t2}a_{j2} + \dots + a_{ti}a_{ji} + \dots + a_{tk}a_{jk} = r_{jt}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_{t1}a_{r1} + a_{t2}a_{r2} + \dots + a_{ti}a_{ri} + \dots + a_{tk}a_{rk} = r_{rt}$$

Applying the theory of least squares we get the following normal equations:

$$a_{t1}\sum_j a_{j1}^2 + a_{t2}\sum_j a_{j1}a_{j2} + \dots + a_{ti}\sum_j a_{j1}a_{ji} + \dots + a_{tk}\sum_j a_{j1}a_{jk} = \sum_j a_{j1}r_{jt}$$

(1)

$$a_{t1}\sum_j a_{j1}a_{j2} + a_{t2}\sum_j a_{j2}^2 + \dots + a_{ti}\sum_j a_{j2}a_{ji} + \dots + a_{tk}\sum_j a_{j2}a_{jk} = \sum_j a_{j2}r_{jt}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_{t1}\sum_j a_{j1}a_{jk} + a_{t2}\sum_j a_{j2}a_{jk} + \dots + a_{ti}\sum_j a_{ji}a_{jk} + \dots + a_{tk}\sum_j a_{jk}^2 = \sum_j a_{jk}r_{jt}$$

The solution of these k normal equations gives the desired projections, a_{t1}, a_{t2}, ..., a_{ti}, ..., a_{tk}. The communality for test t is then: $h_t^2 = a_{t1}^2 + a_{t2}^2 + \dots + a_{ti}^2 + \dots + a_{tk}^2$. This method is applicable to any orthogonal system but is especially adapted to the principal component method of factoring where $\sum_j a_{ji}\, a_{jk} = 0$ for $i \neq k$.

Rewriting the normal equations (1) in matrix form:

$$\begin{bmatrix} \sum_j a_{j1}^2 & \sum_j a_{j1}a_{j2} & \cdots & \sum_j a_{j1}a_{ji} & \cdots & \sum_j a_{j1}a_{jk} \\ \sum_j a_{j1}a_{j2} & \sum_j a_{j2}^2 & \cdots & \sum_j a_{j2}a_{ji} & \cdots & \sum_j a_{j2}a_{jk} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_j a_{j1}a_{ji} & \sum_j a_{j2}a_{ji} & \cdots & \sum_j a_{ji}^2 & \cdots & \sum_j a_{ji}a_{jk} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_j a_{j1}a_{jk} & \sum_j a_{j2}a_{jk} & \cdots & \sum_j a_{ji}a_{jk} & \cdots & \sum_j a_{jk}^2 \end{bmatrix} \begin{bmatrix} a_{t1} \\ a_{t2} \\ \cdots \\ a_{ti} \\ \cdots \\ a_{tk} \end{bmatrix} = \begin{bmatrix} \sum_j a_{j1}r_{jt} \\ \sum_j a_{j2}r_{jt} \\ \cdots \\ \sum_j a_{ji}r_{jt} \\ \cdots \\ \sum_j a_{jk}r_{jt} \end{bmatrix}$$

$$(2) \qquad A \qquad T = P$$

Let the left hand matrix be denoted as A, the middle matrix as T, the right hand matrix as P. Matrix A in this problem is of order 16 x 16. It is a symmetric matrix. Matrix T is the matrix of unknowns and its order is 16 x 11. The order of matrix P is also 16 x 11. All elements in matrices A and P can be determined.

The orthogonal factor matrix from the reasoning battery is reproduced here in Table 2. This factor matrix when pre-multiplied by its transpose gives matrix A. Matrix A appears in Table 3.

The correlations between each of the eleven added tests and the original sixty-six tests appear in Table 4. This matrix of correlations when pre-multiplied by the transpose of the orthogonal factor matrix gives the product matrix P. This product matrix is given in Table 5.

If both sides of the matrix equation (2) are pre-multiplied by the inverse of matrix A, we get:

$$A^{-1}A\ T = A^{-1}\ P$$

but

$$A^{-1}A = I \text{ (the identity matrix)}$$

therefore

$$T = A^{-1}\ P.$$

The inverse of matrix A is given in Table T. When the rows of A^{-1} are multiplied by the columns of P, we get the projections of the added tests on the sixteen dimensions of the reasoning battery. Table 7 gives these projections and the communalities for each of the eleven added tests.

Table 8 gives the transformation matrix obtained by Dr. Adkins when the reasoning battery was rotated to simple structure. Table 9 shows the projections of these added variables upon the oblique reference axes obtained by Dr. Adkins.

The projections of the eleven added tests, shown in Table 7, are consistently low for several of the tests. This seems to indicate that the factorial composition of these tests with low communality cannot be adequately described in terms of the factors isolated in the reasoning study. Without a complete refactorization of all seventy-seven tests it is not possible to say whether or not some additional factors would be generated by the addition of these eleven variables.

Table 1

REASONING STUDY TEST BATTERY

Test No.	Test Name	Test No.	Test Name
1.	Absurdities	45	Progressive Matrices B
2.	Arithmetic	46.	Progressive Matrices C
3.	Block Counting	47.	Progressive Matrices D
4.	Camouflaged Outlines	48.	Progressive Matrices E
5.	Cards	49.	Reading
6.	Circles	50.	Reading II
7.	Conclusions	51.	Reasons
8.	Decoding	52.	Sentence Order
9.	Designs	53.	Series
10.	False Premises	54.	Street Gestalt Completion
11.	Figure Analogies	55.	Suffixes
12.	Figure Classification I	56.	Surface Development
13.	Figure Classification IIA	57.	Things Round
14.	Figure Classification IIB	58.	Topics
15.	Figure Series	59.	Verbal Analogies
16.	Figures	60.	Verbal Classification I
17.	First and Last Letters	61.	Verbal Classification II
18.	Forms	62.	Vocabulary
19.	Geometrical Puzzles	63.	Word-Group Naming
20.	Identical Forms	64.	Word Selection
21.	Incomplete Outlines	65.	Word Squares
22.	Letter Series	66.	Education
23.	Logical Puzzles	*	FOLLOWING TEST TO BE ADDED
24.	Map Planning	67.	Reading and Vocabulary Test (RV-1)
25.	Matrices VI		
26.	Mechanical Information	68.	Arithmetic Reasoning Test (AR-1)
27.	Mechanical Movements		
28.	Mixed Series	69.	Pattern Analysis Test (PA-1)
29.	Mutilated Pictures		
30.	Mutilated Words	70.	Mechanical Aptitude Test (MA-5)
31.	Nim		
32.	Number Series	71.	Army Clerical Speed Test (ACS-1)
33.	Numerical Operations I		
34.	Numerical Operations II	72.	Army Radio Code Aptitude Test (ARC-1)
35.	Numerical Puzzles		
36.	Overlapping Circles	73.	Shop Mechanics Test (SM-1)
37.	Painted Blocks		
38.	Paper Folding	74.	Automotive Information Test (AI-1)
39.	Picture Analogies		
40.	Picture Arrangement	75.	Electrical Information Test (EI-1)
41.	Picture Classification		
42.	Picture-Group Naming	76.	Radio Information Test (RI-1)
43.	Planning a Circuit		
44.	Practical Situations	77.	Criterion Rating

The analysis of the reasoning battery included tests 1-66. Tests 67-77 inclusive are the tests which are to be added.

Table 2

CENTROID FACTOR MATRIX F_0

Test No.	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
1	56	-40	-12	27	-07	-12	08	-12	08	-07	15	-06	-07	-04	-07	-04
2	65	03	-14	-07	-17	-07	-07	-12	20	-08	-06	05	-11	11	11	-08
3	58	26	31	05	-09	-10	09	-07	-03	13	02	-12	06	03	-17	12
4	54	18	14	18	11	-03	06	16	-15	03	05	-14	-05	-15	10	16
5	56	37	18	19	-22	10	04	-06	09	-20	-12	-06	10	13	-07	08
6	58	18	04	-14	-06	-15	05	02	-07	03	04	14	02	09	13	-13
7	37	-13	-19	-10	07	12	06	-07	-16	03	-14	02	-05	04	-08	-05
8	71	09	-08	02	13	-11	-02	17	-02	-02	05	16	-09	-04	08	-09
9	54	16	13	24	-06	04	-10	-06	-15	-07	02	12	-06	-09	-06	05
10	38	06	-20	-15	08	11	17	-17	-03	-18	16	-17	07	17	09	-10
11	70	20	-19	-03	12	17	-02	04	-11	-08	11	-11	-06	08	02	16
12	55	16	-12	-12	04	-02	-06	-06	-12	-06	-18	-07	-13	02	-07	-16
13	60	31	-10	09	-11	-25	-03	20	-10	13	-10	-11	14	-13	-02	-02
14	62	31	-20	-10	-05	-09	09	09	-21	14	-06	-07	10	-09	07	-05
15	67	16	17	-12	20	-10	-02	06	04	11	-05	14	12	14	06	05
16	53	33	18	14	-10	11	11	03	16	-19	-17	10	11	11	-13	08
17	44	-20	-15	-13	-24	23	-07	12	-12	14	05	-11	14	-04	-11	07
18	48	-03	10	-11	06	-07	-17	10	19	16	13	-18	-08	19	-10	-10
19	60	24	11	05	-07	-05	05	12	-05	-07	-10	-15	05	05	-06	-05
20	52	-10	39	14	12	-12	03	-14	-22	19	08	-16	12	-07	12	-09
21	80	-10	-12	-08	14	-06	-08	-06	-10	05	09	-02	08	06	-05	-05
22	69	-01	03	-06	02	-16	-02	-04	07	07	-07	14	-03	-03	10	09
23	65	-19	-13	08	-12	-02	14	-10	16	-03	05	06	04	-07	17	01
24	58	14	20	06	-06	-04	12	18	-03	10	-08	14	13	-06	16	-10
25	78	16	20	-05	11	-01	03	06	18	-03	-08	01	07	-09	02	10
26	33	16	-14	33	-28	09	-13	05	-05	13	26	13	-06	13	02	08
27	40	23	-10	20	-13	13	-20	18	03	11	12	-12	-10	04	13	-09
28	80	-02	-05	-03	09	03	-06	-06	-12	06	-13	08	21	12	05	-04
29	42	-04	37	12	25	10	03	-16	-11	08	06	07	-10	-07	07	14
30	32	-18	16	19	08	16	-28	-07	05	04	-07	08	06	-14	-05	-14
31	32	11	04	-08	-11	-14	-15	09	17	-10	03	05	-07	-11	-12	10
32	63	15	-12	-21	-07	-10	-04	-18	07	-04	-11	07	03	05	07	-12
33	57	-26	09	-24	-18	-14	-15	-24	-05	-16	13	17	10	-13	-09	14

Decimals have been omitted.

Table 2 (continued)

CENTROID FACTOR MATRIX F_0

Test No.	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
34	69	-12	12	-22	-16	-19	-13	-18	-09	-10	11	19	19	-08	02	02
35	34	-13	-03	05	-25	06	04	-04	03	-06	04	-04	09	11	-02	-06
36	52	25	08	-05	12	-14	07	09	04	-02	12	-02	-06	-12	-19	-14
37	55	-11	-05	-04	-28	-13	-14	-13	13	17	06	-11	-14	08	04	-06
38	62	35	08	08	-04	07	-11	-12	-05	-10	08	-06	-05	04	07	-06
39	55	07	-08	-17	07	-08	-07	-08	-04	-24	-10	-16	-13	-13	-08	03
40	58	-12	16	11	26	-08	08	-27	14	-06	-05	05	02	04	-06	-03
41	40	-11	15	14	-08	-17	13	-13	15	13	-06	-13	-18	-08	08	03
42	53	-33	13	03	09	17	-08	09	13	-05	-20	-11	-12	02	16	08
43	44	29	10	21	-12	09	-10	-13	-14	08	08	02	-07	12	-06	-09
44	47	-02	-26	13	05	16	07	10	16	06	17	14	09	-16	-03	-15
45	54	26	-17	-11	19	13	14	-08	10	12	06	12	-03	-03	-10	08
46	57	20	-19	-19	16	16	07	-14	10	17	-03	09	04	-20	05	08
47	54	31	-18	-18	27	04	11	-07	11	19	-06	-02	07	-08	07	-04
48	65	23	12	-15	11	08	-04	-04	14	12	10	-08	09	05	08	11
49	68	-37	-16	19	10	-08	-10	10	12	-07	07	08	12	-05	-07	-03
50	63	-26	-17	11	04	-13	06	13	03	-11	14	12	06	19	-07	11
51	31	-39	17	-27	-18	10	22	26	-06	17	-05	02	-06	06	-06	-05
52	61	-42	-05	16	09	-19	19	08	06	-16	05	04	11	05	-08	-05
53	64	20	-20	-08	05	16	07	-02	-08	-18	10	04	03	03	-08	-06
54	36	-21	24	16	32	27	-14	09	-04	-08	-14	06	-02	-08	07	-22
55	42	-29	-06	-05	-16	29	-16	15	-12	-11	-17	-12	12	-06	-07	17
56	56	39	27	12	-05	11	-12	-09	-06	06	-07	11	-06	03	-03	09
57	38	-38	16	-16	-17	06	12	26	-15	-10	03	11	-14	-12	12	-14
58	31	-36	19	-30	-23	24	31	08	14	14	-05	04	-22	-08	-17	-10
59	76	-04	-31	-10	08	02	05	04	09	-16	03	02	03	05	19	06
60	77	-14	-10	13	10	-22	04	-03	-15	09	-15	06	-10	13	-07	09
61	78	-09	-16	12	12	-08	07	04	-24	06	-12	07	-13	14	-05	12
62	74	-34	-21	13	06	04	-13	02	12	02	04	-04	03	-06	-08	-03
63	73	-34	02	-09	-06	-04	-03	10	06	-03	-07	-13	03	07	04	12
64	70	-05	-24	10	-08	08	14	-10	09	-16	02	-10	-04	-08	10	06
65	77	-04	02	-05	07	-14	12	06	-15	-08	07	-16	-11	14	11	04
66	56	-20	-06	-27	09	-03	-20	15	-05	-09	02	06	-07	-13	-03	02

Decimals have been omitted.

Table 3

$F_o' F_o = A$

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
I	22.1345	2030	-3910	-0379	3881	-4665	0321	-0738	-0270	-1767	0170	1378	2657	1910	2318	1910
II	2030	3.4334	1443	1498	1264	0137	-0644	-2508	-1340	2087	-0369	-0452	1239	1302	0775	0844
III	-3910	1443	1.8234	1938	-0263	-0387	-0316	-0487	-0791	1832	-2005	-0040	-0510	-0857	-0291	0378
IV	-0379	1498	1938	1.4582	0168	0014	-1405	-0208	-0289	-0231	1068	-0565	-0109	0857	0002	0385
V	3881	1264	-0263	0168	1.3281	-0116	0538	-0947	-0240	0357	-0914	0475	0382	-0824	0950	-0186
VI	-4665	0137	-0387	0014	-0116	1.0881	-0405	1026	0202	-0098	-0699	-0620	-0273	-0585	-0258	-0108
VII	0321	-0644	-0316	-1405	0538	-0405	.8775	0467	0467	0346	-0340	0298	-0338	0355	0227	-0156
VIII	-0738	-2508	-0387	-0208	-0947	1026	0467	.9658	-1001	0752	-0510	-0367	-0300	-0506	-0063	-0430
IX	-0270	-1340	-0791	-0289	-0240	0202	0467	-1001	.9013	-0545	0325	0348	-0160	-0374	-0083	-0195
X	-1767	2087	1832	-0231	0357	-0098	0346	0752	-0545	.8581	0225	-0311	-0388	-0401	0478	-0524
XI	0170	-0369	-2005	1068	-0914	-0699	-0340	-0510	0325	0225	.6779	0345	0124	0210	0020	-0262
XII	1378	-0452	-0040	-0565	0475	-0620	0298	-0367	0348	-0311	0345	.7124	0798	-0549	-0065	-0088
XIII	2657	1239	-0510	-0109	0382	-0273	-0338	-0300	-0160	-0388	0124	0798	.6186	-0146	-0082	0359
XIV	1910	1302	-0857	0857	-0824	-0585	0355	-0506	-0274	-0401	0210	-0549	-0146	.6416	-0023	0229
XV	2318	0775	-0291	0002	0950	-0258	0227	-0063	-0083	0478	0020	-0065	-0082	-0023	.5638	-0382
XVI	1910	0844	0378	0385	-0186	-0108	-0156	-0430	-0195	-0524	-0262	-0088	0359	0229	-0382	.5935

Decimals have been omitted except for diagonal elements.

Table 4

CORRELATIONS BETWEEN ADDED TESTS AND REASONING BATTERY TESTS

Test No.	67	68	69	70	71	72	73	74	75	76	77
1	51	47	23	39	36	29	39	30	32	07	26
2	53	71	43	41	31	22	38	26	32	12	19
3	25	40	56	43	37	20	34	28	27	01	23
4	28	30	50	33	16	16	33	25	35	05	25
5	25	46	57	43	28	32	39	35	39	06	20
6	34	47	39	36	29	22	37	22	18	-04	14
7	30	25	20	16	26	16	17	-03	13	02	07
8	50	60	51	36	40	32	44	36	33	07	27
9	31	38	49	42	24	23	39	37	37	04	21
10	31	33	30	12	27	18	09	04	25	14	17
11	44	45	49	38	31	36	35	28	36	12	32
12	40	45	35	26	33	25	24	15	25	05	16
13	40	49	52	40	18	27	40	33	35	07	25
14	39	46	47	33	30	24	32	24	37	09	24
15	41	48	53	36	35	36	34	21	22	01	21
16	19	33	48	32	20	24	23	20	23	04	13
17	37	32	24	26	24	24	34	16	22	02	21
18	30	32	30	32	30	23	24	09	25	01	21
19	36	44	53	47	25	26	33	18	32	05	11
20	30	35	41	35	40	22	34	20	21	07	28
21	65	61	56	43	49	31	47	24	42	06	29
22	47	58	40	40	44	30	47	34	32	19	25
23	50	57	33	43	30	27	40	26	24	20	21
24	36	46	51	42	29	23	41	23	27	01	13
25	47	59	61	43	42	28	45	25	30	12	29
26	17	29	25	41	01	21	42	71	29	08	27
27	26	35	28	37	14	19	33	40	32	05	18
28	60	64	53	42	45	31	41	28	32	06	26
29	20	22	38	26	35	17	15	19	21	04	26
30	29	27	24	23	29	14	27	24	19	07	15
31	28	30	24	24	17	20	28	15	13	08	13
32	47	57	43	28	43	19	35	23	28	13	15
33	46	51	32	38	58	34	35	17	23	16	23

Test No.	67	68	69	70	71	72	73	74	75	76	77
34	51	61	44	39	55	38	36	24	27	17	22
35	15	32	19	24	21	17	23	14	26	09	09
36	32	37	42	29	37	27	33	17	18	01	18
37	40	52	26	45	27	24	32	27	33	06	18
38	31	49	60	41	20	29	33	29	34	10	18
39	46	44	35	29	37	25	31	07	30	16	15
40	46	43	45	32	37	23	31	09	21	03	10
41	27	30	16	27	21	17	25	11	11	02	03
42	47	42	27	37	32	27	35	22	22	12	22
43	23	30	47	36	24	27	30	34	28	01	14
44	31	38	28	29	24	24	35	30	20	00	31
45	26	36	46	27	27	19	30	20	20	01	21
46	37	44	45	25	32	18	30	18	25	07	21
47	30	40	47	21	29	21	27	17	20	05	12
48	35	51	56	36	39	27	30	18	30	06	26
49	67	52	32	45	39	37	50	32	31	09	27
50	58	46	35	42	26	36	36	28	24	10	25
51	27	15	04	15	21	13	10	01	09	04	12
52	60	46	27	41	41	33	42	20	21	00	23
53	42	45	51	32	34	31	36	22	30	03	24
54	28	18	23	25	29	15	19	12	16	15	07
55	38	32	21	27	25	24	28	14	20	11	15
56	23	42	68	42	29	29	38	41	26	-02	25
57	37	26	10	16	32	21	16	04	08	02	11
58	16	20	07	12	28	07	08	-03	01	05	13
59	61	64	42	42	39	34	45	27	34	06	31
60	63	56	47	49	42	37	47	31	34	00	32
61	64	59	52	47	41	40	47	31	39	05	35
62	72	62	37	47	40	34	54	29	38	10	30
63	58	54	35	46	44	32	47	25	41	12	25
64	56	60	46	47	35	33	48	35	37	19	27
65	50	55	48	49	38	38	45	24	41	12	27
66	52	50	27	32	46	34	39	15	25	12	34

Decimals have been omitted.

Table 5

PRODUCT MATRIX $P = F'_o R_a$

R_a = Correlation Matrix of Added Tests vs Reasoning Battery Tests

Dimension No.	Added Tests 67	68	69	70	71	72	73	74	75	76	77
I	15.7255	17.1918	15.2227	13.5197	12.4217	10.0895	13.2731	8.7890	10.4592	2.6978	8.1468
II	-8709	1008	1.4501	1406	-4515	-0629	0335	5520	3600	-1842	-0319
III	-7056	-5322	1086	-0692	0218	-2186	-3326	-2140	-2962	-1161	-2270
IV	-0889	-1002	2073	3964	-3510	0625	3091	6848	2448	-0793	0925
V	3290	0982	4098	-0313	3148	1011	0684	-1686	0362	-0348	1555
VI	-4581	-4827	-2244	-3228	-2891	-2259	-3278	-1064	-1633	-0048	-0926
VII	-0814	-1032	-0379	-1406	-0284	-0848	-1644	-2747	-1556	-0801	-0647
VIII	0437	-1751	-2276	-0038	-2113	0280	0464	0081	-0407	-0752	0621
IX	0139	1269	-1483	-0045	-0334	-0657	0221	-0406	-1113	0514	-0784
X	-3103	-2228	-0371	-0851	-1622	-1979	-0939	0570	-1129	-1731	0035
XI	0059	0364	-0206	0946	0009	0906	1018	2385	0848	0074	1810
XII	1163	1544	0918	0359	1547	0926	1124	1670	-1085	-0333	0522
XIII	1724	2222	3013	1249	1653	1260	1746	0784	1086	0292	0753
XIV	0877	1445	1632	1779	-0367	1300	0449	1240	1548	-0568	0267
XV	1486	2592	1359	1240	0430	0368	1234	1500	1285	1063	0537
XVI	0754	0937	2084	1677	0231	1298	1488	1572	1284	0567	1681

Decimals have been omitted except for those values greater than unity.

Table 6

A^{-1}

Factor No.	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
I	0467	-0017	0086	0012	-0125	0186	0004	-0023	0005	0077	0013	-0059	-0160	-0126	-0175	-0146
II	-0017	3140	-0124	-0249	-0201	-0181	0163	0851	0424	-0856	0218	0164	-0649	-0631	-0361	-0370
III	0086	-0124	6099	-0999	0369	0380	0095	0554	0430	-1355	2071	-0196	0456	0941	0354	-0330
IV	0012	-0249	-0999	7335	-0355	-0105	1095	-0191	0051	0461	-1511	0504	0150	-1058	-0070	-0369
V	-0125	-0201	0369	-0355	7976	0018	-0554	0924	0306	-0362	1280	-0443	-0347	1192	-1166	0299
VI	0186	-0181	0380	-0105	0018	9571	0463	-0964	-0353	0189	0999	0724	0295	0850	0370	0099
VII	0004	0163	0095	1095	-0554	0463	1.1846	-0666	-0693	-0490	0481	-0519	0638	-1015	-0355	0078
VIII	-0023	0851	0554	-0191	0924	-0964	-0666	1.1207	1413	-1179	1022	0391	0166	0854	0043	0622
IX	0005	0424	0430	0051	0306	-0353	-0693	1413	1.1497	0478	-0338	-0425	0330	0611	0094	0356
X	0077	-0856	-1355	0460	-0362	0189	-0490	-1179	0478	1.2561	-1123	0520	0683	0677	-0875	1007
XI	0013	0218	2071	-1511	1280	0999	0481	1022	-0338	-1123	1.6115	-0843	-0196	0026	-0040	0673
XII	-0059	0164	-0196	0504	-0443	0724	-0519	0391	-0425	0520	-0843	1.4616	-1796	1202	0243	0279
XIII	-0160	-0649	0456	0150	-0347	0295	0638	0166	0330	0683	-0169	-1796	1.6835	0504	0342	-0852
XIV	-0126	-0631	0941	-1058	1192	0850	-1015	0854	0611	0677	0026	1202	0504	1.6567	0094	-0338
XV	-0175	-0361	0354	-0070	-1166	0372	-0355	0043	0094	-0875	-0040	0243	0342	0094	1.8266	1133
XVI	-0146	-0370	-0330	-0369	0299	0099	0078	0622	0356	1007	0673	0279	-0852	-0338	1133	1.7324

Decimals have been omitted except for those values greater than unity.

Table 7

ADDED SECTION OF ORTHOGONAL FACTOR MATRIX

Test	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	h^2
67	706	-281	-219	009	065	-142	-157	041	-012	-108	-099	-001	-005	-041	-005	-061	703220
68	774	-007	-123	-033	-174	-115	-159	-115	140	-056	-061	057	010	-035	167	-090	764370
69	688	354	172	095	074	104	-004	-085	-083	-010	024	016	124	-002	-086	046	693515
70	615	-004	069	259	-186	-041	-135	051	038	031	071	-024	-044	050	008	066	524064
71	566	-170	172	-235	083	000	-076	-215	-071	-067	-058	046	039	-138	-142	-140	568294
72	456	-038	000	035	-024	-011	-103	075	-062	-122	115	049	-006	065	-094	062	278831
73	601	-021	-085	235	-121	-065	-175	094	053	048	046	059	014	-144	-003	062	517914
74	400	123	-091	453	-253	081	-245	065	-002	165	205	248	-092	-011	132	129	696847
75	474	068	-078	153	-104	040	-147	-005	-093	-035	068	-206	-026	026	046	040	354845
76	122	-054	-018	-052	-070	043	-089	-082	034	-161	006	-087	-002	-134	178	061	130269
77	370	-034	-037	055	038	078	-070	085	-074	090	262	013	-047	-064	-050	195	291322

Table 8

TRANSFORMATION MATRIX FOR THE REASONING BATTERY

Oblique Reference Axes

	A'	B'	C'	D'	E'	F'	G'	H'	I'	J'	K'	L'	M'	N'	O'	P'
I	21	22	19	18	02	19	14	16	-13	21	10	17	22	17	14	18
II	-03	13	-26	06	-17	25	17	-06	-06	-02	-19	25	-06	-14	-09	07
III	-19	-35	-19	28	05	09	02	04	17	16	-07	28	07	41	18	-10
IV	07	-30	42	-17	18	03	01	-15	20	-21	-02	09	-16	32	-11	32
V	31	22	32	23	-33	-06	14	16	09	-35	-12	-27	10	37	-08	-11
VI	-04	34	-17	-23	-12	-28	-06	11	09	-37	41	27	00	42	22	20
VII	-01	36	14	02	31	22	-16	23	22	-30	-21	31	-27	-08	40	-38
VIII	-02	-24	17	23	-51	50	16	-12	16	-45	18	-12	29	-06	47	21
IX	-54	36	28	11	17	-20	-08	-19	-01	06	-08	32	43	-12	00	14
X	-22	31	06	29	41	14	-39	-19	-12	00	08	-31	00	15	-02	20
XI	-17	12	20	37	-05	06	10	36	19	06	-10	-36	-11	04	16	61
XII	09	09	09	-60	-40	-21	-13	-29	20	41	-48	-06	-18	-11	35	24
XIII	-60	25	26	05	04	30	-52	42	10	18	60	17	-25	08	-25	-29
XIV	15	-23	00	05	06	-38	-57	28	-39	-14	-21	18	05	-36	-07	-08
XV	-05	08	σ49	σ29	04	37	-25	46	08	27	03	-37	66	25	-22	-10
XVI	24	-13	-25	11	-29	-21	-19	-31	75	-20	18	-17	15	-34	-48	-16

Decimal points have been omitted.

Table 9

OBLIQUE FACTOR MATRIX

Reference Axes

Test	A'	B'	C'	D'	E'	F'	G'	H'	I'	J'	K'	L'	M'	N'	O'	P'
67	26	06	27	03	-05	07	17	06	-18	18	11	-05	21	06	02	10
68	05	18	05	-05	10	13	08	11	-24	45	04	10	31	03	-02	16
69	11	17	08	17	-05	17	11	13	-01	11	08	29	-01	23	04	13
70	09	-08	14	12	07	11	07	-01	-04	15	07	10	17	10	02	31
71	08	17	11	14	03	-02	19	15	-15	32	04	09	-03	23	14	03
72	17	-05	14	09	-14	02	12	05	00	07	04	04	05	00	08	19
73	07	03	22	05	-02	17	15	-11	03	18	12	-01	16	11	02	33
74	08	-06	09	-14	-05	07	01	-16	13	18	-01	-11	12	11	00	60
75	15	-04	02	12	05	12	12	11	-09	03	18	03	13	10	-10	19
76	00	02	-15	-06	-01	03	09	09	06	13	12	-03	18	06	-12	-02
77	14	07	11	19	-12	05	10	01	17	-03	10	-15	06	09	05	30

Decimal points have been omitted.

3. Estimation of Reliability of Mental Tests from Parallel Forms.

The reliability coefficient of a test may be defined to be the correlation coefficient between two parallel forms of a test. In the development of this fundamental concept in the theory of mental tests the notion of the existence of an indefinitely large number of parallel forms of a particular test is basic.

If we assume that a test score is of the form

$$X_{ig} = T_i + E_{ig}$$

where g is a subscript denoting the test and i the subscript denoting the individual, then the population parameter to be estimated is

$$\rho_{xx} = \frac{\sigma_{gh}}{\sigma_g \sigma_h}$$

where h is a subscript denoting a test which is a parallel form of test g.

Suppose we select a random sample of K tests from the group of parallel forms and apply these tests to each of N individuals who are chosen in some specified way. Each application of a parallel form results in N scores. Consequently the matrix of observations is as follows.

		Tests				
		1	2	3	...	K
	1	X_{11}	X_{12}	X_{13}	...	X_{1K}
	2	X_{21}	X_{22}	X_{23}	...	X_{2K}
Individuals	3	X_{31}	X_{32}	X_{33}	...	X_{3K}
	.	.	.	.	...	.
	.					
	N	X_{N1}	X_{N2}	X_{N3}		X_{NK}

This set of scores may be regarded as arising from a K variate probability distribution where $E(X_{i1}) = E(X_{i2}) = \dots = E(X_{iK}) = \mu$, (E denotes "expected" value) the variance $\sigma^2_{X_g} = \sigma^2$ for all g and $\rho_{gh} = \rho$ for all g and h. These are necessary conditions for a set of tests to be regarded as parallel forms.

The parameter ρ is the intra-class correlation coefficient or, in terms of our previous discussion, the reliability coefficient.

In order to obtain an estimate of ρ we use the likelihood ratio criterion and obtain a confidence interval procedure.

We transform the original scores by means of an orthogonal transformation so that

$$Y_{i1} = \frac{X_{i1} + X_{i2} + \dots + X_{iK}}{K^{1/2}} = \sqrt{K}\ \bar{X}_i$$

and

$$Y_{ig} = c_{g1}X_{i1} + c_{g2}X_{i2} + \dots + c_{gK}X_{iK}$$

where

$$\sum_{j=1}^{K} c_{gj} = 0 \; ; \qquad \sum_{j=1}^{K} c_{gj}^2 = 1 \; ;$$

and

$$\sum_{j=1}^{K} c_{gj} c_{hj} = 0 \qquad \text{for } g \neq h \; .$$

As a result of this transformation we have a transformed observation matrix

$$\begin{matrix} Y_{11} & Y_{12} & \cdots & Y_{1K} \\ Y_{21} & Y_{22} & \cdots & Y_{2K} \\ \cdot & \cdot & \cdot & \cdot \\ Y_{N1} & Y_{N2} & \cdots & Y_{NK} \end{matrix}$$

where

$$E(Y_{i1}) = \sqrt{K}\;\mu, \; E(Y_{ij}) = 0 \quad j \neq 1 \quad \sigma_{Y_1}^2 = \sigma^2[1 + (K-1)\,\rho] = \sigma_1^2$$

$$\sigma_{Y_j}^2 = \sigma^2\,[1 - \rho] = \sigma_2^2 \quad \text{and} \quad \rho_{gh} = 0.$$

We may observe at this point that the expression σ_1^2 is necessarily not less than zero. This implies $\sigma^2[1 + (K-1)\,\rho] \geq 0$, hence $\rho \geq \frac{-1}{K-1}$ which shows that the value of the reliability coefficient is directly affected by the number of parallel forms used to estimate it. We shall return to this point and others logically connected with it later.

The frequency function for the sample of Y's may now be written as

$$\frac{1}{(2\pi)^{\frac{1}{2}N}\ \sigma_1^N}\ e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{N}(Y_{i1}-\sqrt{K}\ \mu)^2}\qquad \frac{1}{(2\pi)^{\frac{1}{2}N(K-1)}\ \sigma_2^{N(K-1)}}\ e^{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{N}\sum_{g=2}^{K}(Y_{ig})^2}\ .$$

It follows that

$$U = \sum_{i=1}^{N}(Y_{i1} - \bar{Y}_1)^2 = \sigma_1^2\chi^2_{N-1}$$

and

$$V = \sum_{i=1}^{N}\sum_{g=2}^{K} Y_{ig}^2 = \sigma_2^2\ \chi^2_{(K-1)N}$$

where

$$\bar{Y}_1 = \sum_{i=1}^{N}\frac{Y_{i1}}{N}\ ;\qquad \chi^2_{N-1} = \chi^2_1 \text{ and } \chi^2_{(K-1)N} = \chi^2_2$$

are chi-square random variables and independent.

Our purpose is to estimate ρ and to test hypotheses covering it. Therefore we consider the ratio

$$\frac{\sigma_1^2}{\sigma_2^2} = \lambda = \frac{1 + (K - 1)\rho}{1 - \rho}$$

which is independent of unknown parameters and is in one to one correspondence with ρ. The likelihood ratio statistic λ is estimated as

$$\frac{U}{V} = \lambda \frac{\chi_1^2}{\chi_2^2}$$

and the ratio U/V has the well-known Beta frequency function

$$\frac{1}{B\left[\frac{N-1}{2}, \frac{N(K-1)}{2}\right]} \frac{1}{\lambda} \frac{\left(\frac{\chi}{\lambda}\right)^{\frac{N-1}{2} - 1}}{\left[1 + \frac{\chi}{\lambda}\right]^{\frac{1}{2}NK}} .$$

Although in most practical problems there is no interest in testing hypotheses concerning ρ the test is derived here to round out the discussion.

To test the hypothesis H_0: $\lambda = \lambda_0$ against the alternative

$$H : \lambda \neq \lambda_0$$

we require a rule for rejection of the hypothesis H_0 if $\frac{U}{V} \leq c_1$ or if $\frac{U}{V} \geq c_2$ where c_1 and c_2 are constants determined by

$$\text{Probability} \left[\frac{U}{V} \leq c_1 \text{ or } \frac{U}{V} \geq c_2 \text{ for } H_0 \text{ true} \right] = \epsilon$$

and

$$\frac{d}{d\lambda} \text{Probability} \left[\frac{U}{V} \leq c_1 \text{ or } \frac{U}{V} \geq c_2 \text{ given that H is true} \right]_{\lambda=\lambda_0} = 0.$$

Letting $\frac{c_1}{\lambda} = \gamma_1$ and $\frac{c_2}{\lambda} = \gamma_2$ we may determine a best test by solving

$$(1) \qquad \frac{1}{B\left[\frac{(N-1)}{2}, \frac{N(K-1)}{2}\right]} \int_{\gamma_1}^{\gamma_2} \frac{(X)^{\frac{N-1}{2}-1}}{\left[1+X\right]^{\frac{1}{2}NK}} \, dX = 1 - \epsilon$$

and

$$(2) \qquad \frac{\gamma_2^{\frac{N-1}{2}}}{\left[1+\gamma_2\right]^{\frac{1}{2}NK}} = \frac{\gamma_1^{\frac{N-1}{2}}}{\left[1+\gamma_1\right]^{\frac{1}{2}NK}}$$

where ϵ is the probability of making the error of rejecting H_0 when it is in fact true. The two equations in γ_1 and γ_2 determine the rule for rejection of the hypothesis H_0 at the ϵ level of significance. Values of γ_1 and γ_2 are tabulated for given values of N and K.

For example if H_0: $\rho = 0$ substitute and get $\lambda_0 = 1$ then do not reject H_0 if

$$\gamma_1 \leq \frac{U}{V} \leq \gamma_2$$

In general the rule for non-rejection is

$$\gamma_1\lambda \leq \frac{U}{V} \leq \gamma_2\lambda \, .$$

In order to estimate ρ which is usually what is wanted the statement above is converted as

$$\frac{1}{\gamma_2}\frac{U}{V} \leq \lambda \leq \frac{U}{V}\frac{1}{\gamma_1}$$

which is a true statement with probability $1 - \epsilon$.

If $N \geq 10$ the equal tails of the F distribution may be used to determine γ_1 and γ_2. In this case the probability level is not ϵ but slightly greater. The difference is not, however, of practical importance. We determine γ_1 and γ_2 as

$$\Pr\left[F_{N-1,N(K-1)} \geq \gamma_2\right] = \frac{\epsilon}{2}$$

and

$$\Pr\left[F_{N-1,N(K-1)} \leq \gamma_1\right] = \frac{\epsilon}{2}$$

where γ_1 and γ_2 are tabled for conventional values of ϵ . Since many F tables give only critical values for the upper tail of the distribution it should be recalled that

$$F_{N-1,N(K-1),\,\epsilon} = \frac{1}{F_{N(K-1),N-1,1-\epsilon}} \quad .$$

As an example if $K = 2$, $N = 121$, $\epsilon = 10$ o/o, we have to determine

$$\Pr\left[F_{120,120} \geq \gamma_2\right] = .05$$

and

$$\Pr\left[F_{120,120} \leq \gamma_1\right] = .05 = 1 - \Pr\left[F_{120,120} \geq \gamma_1\right] = .95$$

From the ordinary F table $\gamma_2 = 1.352$ and $\gamma_1 = \frac{1}{1.352} = .7396$. If, for example, $K = 3$, $N = 121$, $\epsilon = 10$ o/o, then we have to determine

$$\Pr\left[F_{120,240} \geq \gamma_2\right] = .05$$

$$\Pr\left[F_{120,240} \leq \gamma_1\right] = .05 = 1 - \Pr\left[F_{120,240} \geq \gamma_1\right] = .95 \ .$$

If $N(K-1)$ is large, then $F_{N-1,N(K-1)}$ is approximately distributed like $\chi^2_{N-1/N-1}$. In this case we have

$$\Pr\left[F_{120,240} \geq \gamma_2\right] = \Pr\left[\chi^2_{120} \geq 120\gamma_2\right] = 1.221$$

and

$$\Pr\left[F_{120,240} \leq \gamma_1\right] = 1 - \Pr\left[F_{120,240} \geq \gamma_1\right] \ .$$

Looking in the F table for $\Pr\left[F_{240,120} \geq \frac{1}{\gamma_1}\right] = .05$ we have

$$\frac{1}{\gamma_1} = 1.254 \text{ and } \gamma_1 = .7958.$$

We defined

$$U = \sum_{i=1}^{N} (Y_{i1} - \bar{Y}_1)^2 = \sum_{i=1}^{N} (Y_{i1})^2 - N\,\bar{Y}_1^2 \ .$$

Now

$$Y_{i1} = \frac{X_{i1} + X_{i2} + \dots + X_{iK}}{\sqrt{K}} = \sqrt{K}\ \bar{X}_i \quad .$$

and

$$\bar{Y}_1 = \sqrt{K}\ X..$$

where

$$\bar{X}_i = \sum_{j=1}^{K} \frac{X_{ij}}{K} \quad \text{and} \quad X.. = \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{X_{ij}}{NK}$$

so

$$U = K \sum_{i=1}^{N} \bar{X}_i^2 - NKX.. = K \sum_{i=1}^{N} (\bar{X}_i - X..)^2 \quad .$$

We have

$$V = \sum_{i=1}^{N} \sum_{g=2}^{K} Y_{ig}^2 \quad .$$

Recalling that the X's and Y's are related by an orthogonal transformation we know that

$$\sum_{i=1}^{N} \sum_{h=1}^{K} X_{ih}^2 = \sum_{i=1}^{N} \sum_{h=1}^{K} Y_{ih}^2$$

hence

$$V = \sum_{i=1}^{N} \sum_{h=1}^{K} (X_{ih} - \bar{X}_i)^2 \quad .$$

A compact summary of the calculations described above is as follows. Display the calculations in an analysis of variance table considering the matrix of raw scores as one way classification with regard to individuals. Then

U = Sum of Squares between Individuals

V = Sum of Squares within Individuals.

In such a table we have

Source of Variation	d.f.	SS	M.S.	E.M.S.
Between Individuals	$N - 1$	U	$\frac{S_I}{N-1} = s_I$	$K\sigma_I^2 + \sigma_E^2$
Within Individuals	$N(K - 1)$	V	$\frac{S_E}{N(K-1)} = s_E$	σ_E^2

where $\sigma_I^2 = E \sum_{i=1}^{N} \frac{(T_i - \overline{T})^2}{n-1}$ and E.M.S. means expected values of the square.

With this set up in mind

$$\rho_{xx} = \frac{\sigma_I^2}{\sigma_I^2 + \sigma_E^2}$$

which is estimated as

$$r_I = \frac{s_I - s_E}{s_I - (K-1)s_E}$$

If we have a good estimate of the ratio of σ_I^2 to σ_E^2 and it is equal to α (say), then

$$r_I = \frac{\lambda - 1}{\lambda + (K-1)} \quad .$$

Now if we wish to estimate the value of r_I for some other number of replications, say K', then

$$r_I' = \frac{K r_I}{K' + (K-K')r_I} \quad .$$

If K = pK', then

$$r_I' = \frac{p r_I}{1 + (p-1)r_I}$$

which is the so-called Spearman-Brown prophecy formula for estimating the reliability of a test if it is increased in length p times.

4. Principal Component Factorization

The purpose of factor analysis:

Thurstone* states, "The factorial methods were developed primarily for the purpose of identifying the principal dimensions or categories of mentality; but the methods are general, so that they have been found useful for other psychological problems and in other sciences as well." The factorial method is generally employed by the investigator to test an hypothesis when he is unable to devise some other critical test. Factorial methods may also be used to explore a range of phenomenon to determine the underlying constructs. For example: one might wish to investigate the basic abilities or skills involved in mechanical aptitude. To do this the investigator would design a series of tests which would differentiate between persons with mechanical ability and those without this ability. He would no doubt include tests of manual dexterity, spatial visualization, coordination, etc. Tests of verbal ability would probably be considered not applicable and would not be included in the test battery. This test battery would be given to a group of subjects which has a considerable range in mechanical skill. From the scores obtained on these tests, the investigator would attempt to tease out by factorial methods the underlying abilities involved. His gamble from the beginning of the experiment being that the number of fundamental abilities involved is fewer than the number of tests represented in the test battery.

---

*L. L. Thurstone, "Multiple Factor Analysis," The University of Chicago Press, 1947, Chicago, Illinois, p. 55.

Factor Analysis and Multiple Regression:

In the sense that one is attempting to predict the dependent variable from two or more independent variables, there is no direct relation between factor analysis and multiple regression technique. In the factorial problem we start with the test scores or other measures and attempt to determine the factors. Here the factors could be considered to be the dependent variables and the test scores the independent variables. Generally there is no distinction made between dependent and independent variables in factor analysis. A close parallel exists when one knows the factor scores of an individual and attempts to predict his performance on a test of known factorial composition. Here the factor scores would represent the independent variables while the predicted test score would be the dependent variable.

The fundamental equation of factor analysis is written

$$s_{ji} = c_{j1}x_{1i} + c_{j2}x_{2i} + c_{j3}x_{3i} + \dots + c_{jq}x_{qi} \quad .$$

where s_{ji} is the standard score of individual i in test j, the x's are standard scores of individual i in each of the uncorrelated reference abilities, and the c's are the weights assigned to the standard scores in the reference abilities for the determination of the observed standard score s_{ji}. The equation is written to represent q terms in the right-hand member and it is hoped that this number of reference abilities will be relatively small compared with the number of tests, n, in the whole battery.

The c's are called the test coefficients or factor loadings. To determine these we start with a score matrix S or order n x N, where the elements of S are s_{ji}, the standard score of individual i on test j. N refers to the number of individuals and n denotes the number of tests. The complete correlation matrix is defined as the symmetric matrix of intercorrelations between the tests with unity in the diagonals. This matrix is denoted R_1 and is of order n x n. In matrix notation:

$$R_1 = \frac{1}{N} SS' \text{ with elements}$$

$$r_{jk} = \frac{1}{N} \sum_{i=1}^{N} s_{ji} s_{ki}$$

Most factorial problems are not concerned with explaining the complete variance of the tests. That part of the test variance which is shared with other tests is more commonly of interest. The variance of the test is considered to be divided into three parts, viz., the common variance, the specific variance, and the error variance. The last two are known as the unique variance of the test. The common variance of a test is termed the communality. When the communalities are placed in the diagonal cells, the correlation matrix is known as the reduced correlation matrix and denoted as R.

We wish to find a matrix F, called the orthogonal factor matrix, such that when post multiplied by its transpose will reproduce the correlation matrix R. This factor matrix will have order n x r where r is the number of factors. It is hoped that r will be materially smaller than n, the number of tests.

To obtain the factor matrix F a number of methods have been employed. One of these is the Method of Principal Components set forth by Hotelling.* This method extracts the factors in such a way as to maximize the amount of test variance accounted for by each factor. This method while having desirable characteristics requires considerable computational labor. For desk calculators it represents a staggering amount of work if the test battery is of any size. Consequently other factor methods were devised which would tend to approximate some of the desirable features of the principal component solution but reduced the computational work. The complete centroid method developed by Thurstone is one example of such a compromise. The centroid method does not yield a unique factor matrix. It requires the reflection of test vectors during the extraction of factors and therefore it does not readily lend itself to automatic machine calculation. The principal component method is better suited for machine calculation. Since the fatigue of the operator is not a factor here, the amount of computational labor is not so important as is the case with desk calculators.

Hotelling's method is iterative. It starts with a trial column vector U of n elements. This trial vector is usually taken proportional to the column sums of the reduced correlation matrix. The proportionality factor is so determined as to make the largest entry in U equal to unity. The following matrix multiplication is then carried out

$$R \times U = V$$

*Harold Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, XXIV (September and October, 1933).

This column vector V is again divided by its largest entry to give a second estimate of the vector U. This procedure is repeated until two successive U vectors are obtained which are identical to the required number of decimal places. At this point the test projections on the first factor axis are obtained by multiplying the last column vector V by $1/(U_f'V_f)^{\frac{1}{2}}$, where U_f and V_f are the last values of U and V obtained when convergence has been achieved.

One of the available modifications of the method is that of powering the matrix (i.e., A^t or R^t). The advantages of this procedure together with the mathematics underlying Hotelling's iterative method may readily be seen in the following discussion.

The n characteristic vectors $V_1, \ldots, V_n$ of A, such that

$$V_1 = \begin{bmatrix} V_{11} \\ \vdots \\ V_{n1} \end{bmatrix} \quad \cdots\cdots\cdots\cdots \quad V_n = \begin{bmatrix} V_{1n} \\ \vdots \\ V_{nn} \end{bmatrix} \quad \text{are}$$

determined such that

$$(1) \quad V_i \cdot V_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

or to put it a little more simply $V'V = I$, the unit matrix. It will be assumed that the V_i are so ordered that the associated characteristic roots are $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$. As shown in matrix theory $AV_i = \lambda_i V_i$.

Now any column vector Z of order n x 1 may be expressed in the form

$$(2) \qquad Z = a_1 V_1 + \dots + a_n V_n \;.$$

Thus

$$AZ = a_1 AV_1 + \dots + a_n AV_n = a_1 \lambda_1 V_1 + \dots + a_n \lambda_n V_n$$

$$A^2 Z = a_1 A^2 V_1 + \dots + a_n A^2 V_n = a_1 \lambda^2 V_1 + \dots + a_n \lambda_n^2 V_n$$

. ...

$$(3) \qquad A^t Z = a_1 A^t V_1 + \dots + a_n A^t V_n = a_1 \lambda_1^t V_1 + \dots + a_n \lambda^t V_n$$

Therefore, for the j-th element of $A^t Z$

$$f_j(t) = c_{1j}\lambda_1^t + \dots + c_{nj}\lambda_n^t$$

where $a_i v_{ji} = c_{ij}$.

Then

$$\frac{f_j(t+1)}{f_j(t)} = \frac{c_{1j}\lambda_1^{t+1} + \dots + c_{nj}\lambda_n^{t+1}}{c_{1j}\lambda_1^t + \dots + c_{nj}\lambda_n^t} = \frac{\lambda_1 + \frac{c_{2j}}{c_{1j}}\left(\frac{\lambda_2}{\lambda_1}\right)^t \lambda_2 + \dots + \frac{c_{nj}}{c_{1j}}\left(\frac{\lambda_n}{\lambda_1}\right)^t \lambda_n}{1 + \frac{c_{2j}}{c_{ij}}\left(\frac{\lambda_2}{\lambda_1}\right)^t + \dots + \frac{c_{nj}}{c_{1j}}\left(\frac{\lambda_n}{\lambda_1}\right)^t}$$

and

$$\lim_{t \to \infty} \frac{f_j(t+1)}{f_j(t)} = \lambda_1 \qquad \text{for all } j.$$

Hence, the ratio of every element of the (t+1)st iterate to the t-th iterate approaches λ_1, as t increases and $A^t Z$ approaches V_1 for

$$\frac{A^t Z}{\lambda_1^t} = a_1 V_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^t V_2 + \ldots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^t V_n .$$

Herein the speed of convergence of an arbitrary vector depends on the closeness of the roots and the choice of the initial vector. The further separated the roots are and the closer Z is to the actual vector V_1, the quicker the convergence. It is also evident from the above that the speed depends on the ratio $\frac{\lambda_2}{\lambda_1}$, the closer this quotient is to unity the slower the convergence. If A^t is used, the convergence depends on $\left(\frac{\lambda_2}{\lambda_1}\right)^t$ which is certainly much less than $\frac{\lambda_2}{\lambda_1}$. A word of caution concerning powering should be added; this involves the accumulation of rounding errors as the matrix is raised to high powers, thus making for some loss of efficiency.

Another interesting point is the fact that the diagonal elements of A^t as t is increased tend to be proportional to the squares of the V_{j1}'s, that is to the squares of the elements of the first characteristic vector. The proof is as follows:

In matrix notation (2) may be written

$$Z = Va \text{ where } a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Let $Z_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $Z_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ ... $Z_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$, all n x 1 .

Then since $V'Z = V'Va = a$ from (1)

$$a^{(1)} = \begin{bmatrix} V_{11} \\ \vdots \\ V_{1n} \end{bmatrix} \qquad a^{(2)} = \begin{bmatrix} V_{21} \\ \vdots \\ V_{2n} \end{bmatrix} \quad \cdots \quad a^{(n)} = \begin{bmatrix} V_{n1} \\ \vdots \\ V_{nn} \end{bmatrix}$$

where $Z_i = Va^{(i)}$.

Now the ith elements of A^tZ_i are precisely the diagonal elements of A^t and therefore

$$A^tZ_i = V_{i1}\lambda_1^tV_1 + \ldots + V_{in}\lambda_n^tV_n$$

so that

$$a_{ii}^{(t)} = V_{i1}^2\lambda_1^t + \ . \ . + V_{in}^2\lambda_n^t$$

and

$$a_{jj}^{(t)} = V_{j1}^2\lambda_1^t + \ldots + V_{jn}^2\lambda_n^t$$

where $a_{ii}^{(t)}$ and $a_{jj}^{(t)}$ are the i-th and j-th diagonal elements of A^t.

Dividing $a_{ii}^{(t)}$ by $a_{jj}^{(t)}$,

$$\frac{a_{ii}^{(t)}}{a_{jj}^{(t)}} = \frac{V_{i1}^2\lambda_1^t + \ldots + V_{in}^2\lambda_n^t}{V_{j1}^2\lambda_1^t + \ldots + V_{jn}^2\lambda_n^t} = \frac{V_{i1}^2 + \ldots + V_{in}^2(\lambda_n/\lambda_1)^t}{V_{j1}^2 + \ldots + V_{jn}^2(\lambda_n/\lambda_1)^t}$$

and

$$\lim_{t \to \infty} \frac{a_{ii}^{(t)}}{a_{jj}^{(t)}} = \frac{V_{i1}^2}{V_{j1}^2}$$

This was proved by Sir Cyril Burt in a paper published in the British Journal of Educational Psychology. Much of the material used in this section together with other methods of finding characteristic roots and vectors, is to be found in Mr Seymour Geisser's unpublished master's thesis, University of North Carolina, 1952.

2. Description of the computing procedure:

In adapting this iterative procedure for solution using the IBM calculating punch, the capabilities of the machine have been limited to those present on the standard 602 A machine. Additional features are available on these machines at added cost. If this procedure is to be used on machines possessing additional counter and storage capacity, some modification in the method described here may result in greater efficiency.

The computing method outlined in this paper is a modification of a method reported by King and Priestley.*

A deck of cards is prepared containing the values given in the reduced correlation matrix. Figure 1 shows a sample lay out of matrix cards for a fifteen variable problem. This matrix will have sixteen rows since a sum row is added to provide a summational check for the multiplication of R x U = V. In this fifteen variable problem forty matrix cards will be required. This deck of matrix cards will be used repeatedly from iteration to iteration until the final location of the first factor is determined. Each matrix card is identified by a code number which appears in columns 1-4. This code number identifies the matrix card by row and column. Matrix card no. 1 has the code 0101. It contains elements from the first pair of rows and elements from the first three columns. Matrix card no. 27 would have the code number 0602. Each matrix card contains six elements of the reduced correlation matrix as shown in Figure 1. These data are punched into columns 5-40 inclusive.

A second deck of forty cards will be required for this sample problem. This deck will contain the elements of the vector U Each of these cards, which are called multiplier cards carries three

---

*W. H. King, Jr. and William Priestley, Jr., "Mass Spectrometer Calculations on the IBM 602-A Calculating Punch," IBM Technical Newsletter No. 3, December 1951, pp. 5-17.

elements of the U vector. These multiplier cards are coded also in columns 1-4. The first multiplier card has the code number 0101 and contains elements u_1, u_2, and u_3. The fifth multiplier card has the code number 0105 and contains elements u_{13}, u_{14}, and u_{15}. The five multiplier cards bearing the code 0101, 0102, 0103, 0104, 0105 contain elements of the U vector 1-3, 4-6, 7-9, 10-12, and 13-15 respectively. These multipliers are used in turn to multiply the elements appearing on matrix cards 1-5. Since the storage capacity of the 602 A is not adequate to store these multipliers, it will be necessary to read the values into the machine when multiplying each successive pair of rows. The values from these first five multiplier cards are reproduced seven times so that forty multiplier cards in all are used. The code numbers in columns 1-2 are changed for each reproduction of the multiplier cards so that the first two columns of the card carry code numbers from 01 to 08 in our example.

When both the matrix cards and the multiplier cards are so coded, it is possible to position the correct multiplier card in front of each matrix card by running the cards through the sorter. With the multiplier cards placed face down in the hopper of the sorter and the matrix cards placed face down on top of the multiplier cards, we sort the cards on columns 4, 3, 2, and 1 in that order. The cards are picked up from the stackers from right to left placing the next pack picked up on top of the preceding pack. The cards will then be in the proper order. For this example it is not necessary to sort on all four columns. Only columns 4 and 2 are needed. However, when a large problem is being solved, the code numbers may require two digits and for this reason the space has been provided.

Read Cycle:

The planning chart for this problem is given in Figure 2. The wiring of the control panel appears in Figure 3. The first three elements of vector U which appear in multiplier card 0101 are read into storage units 6L, 7L, and 7R. These multipliers can be either plus or minus. The sign of the value is indicated in columns 41, 47, and 53 for each of the three elements respectively. A 5 punch is used to indicate a negative value while a 0 punch is used to show that the value is positive. Each value is carried to four decimal places. No value can exceed 1.0000. The code number of the multiplier card is read into counter 3. This code number will be used later to indicate which pair of elements of vector V is punched on the trailer card. In order to read the code number in only once, the value is read in through a pilot selector. The eight multiplier cards bearing the code --01 have an 11, sometimes known as an X punch in column 67. This X is sensed by the control brushes and is used to pick up pilot selector six which is impulsed to read and then drop out. The position of control brushes varies from one machine to the next so that the card column corresponding to the position of one of the twenty control brushes in another machine may be different from the ones used here. Column 76 has an 11, or X, punch for all multiplier cards. This punch is used to impulse the skip out hub from the read brushes and also to pick up pilot selector 5 from the control brushes. Pilot selector 5 allows the read hub to be impulsed from the skip out hub for all multiplier cards. It is necessary to impulse read through the pilot selector since for matrix cards we want to go through the program steps but still skip out the matrix card without punching. If the skip out hub were directly wired to the read hub both types of cards would skip out without programming. Pilot selector 5 is wired to read drop out. All multiplier cards have an 11 or X punch in column 78. By means of the control brushes pilot selector 1 is picked up. When transferred pilot selector one reads in the multiplier values into storage units 6L, 7L, and 7R. Pilot selector 1 is wired to read and drop out. It is necessary to use a pilot selector here since we are reading in values from two separate detail cards. If we were to read in the values from the multiplier card directly from the read cycle hubs and then read the matrix card values in from the read cycle hubs, the storage units 6L, 7L, and 7R would be set to zero when the matrix cards were read in. After the values are read in from the first multiplier card this card skips out to the stacker and the first matrix card is read.

The first matrix card carries the first six values from the reduced correlation matrix. These are read into the storage units as indicated in Figure 2 These values can also be either plus or minus. The signs are indicated in columns 5, 11, 17, 23, 29, and 35. Again as for the multiplier cards a 5 punch indicates that the value is negative while as 0 punch shows that the value is positive. These values are punched to three decimal places. Since the column sums will be used to compute a summational check space must be allowed for two whole numbers in addition. Therefore, five columns have been allowed for the numerical value of each entry. The matrix cards all have an 11 or X punch in column 71. This is read by the reading brushes and impulses the skip out hub, however the machine will go through the programs and the card does not skip out until program step 7 is reached.

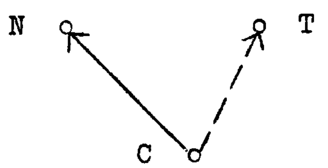
Program Step 1 (For Sign Control)

Since both the multiplier and the multiplicand can be either plus or minus, three pilot selectors must be used to determine the two products developed at any one time. We will multiply h_{11}^2 and r_{21} by u_1 at the same time. The machine will multiply several multiplicands by the same multiplier as rapidly as it will multiply one multiplicand and one multiplier. The efficiency of the machine is increased by developing as many products simultaneously as is possible. In this case only two can be developed at one time. The standard machine has only six counters with a total of thirty positions. Each product for the summational check could have as many as eleven places for a large problem. One additional counter space will be needed to determine the sign of the product. Therefore two twelve position counters will be needed. Counters 1 and 2 are coupled to give one twelve position counter. Counters 4, 5, and 6 are coupled to give a fourteen position counter. Counter 3 is being used for code designation. This exhausts the counter space on the standard 602 A machine. The left hand position of storage unit 6L is read out and wired to permit a 5 to energize pilot selector No 2. If a zero appears in the left hand position, pilot selector No. 2 will not be energized. The remaining five positions of storage unit 6L are read out and read into storage unit 1R. The left hand position of storage unit 2L is read out and wired so that a 5 energizes pilot selector No. 3 while a zero will not. The left hand position of storage unit 3R is read out and wired so that a 5 will energize pilot selector

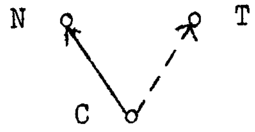
No. 4 while a zero will not. Pilot selector No. 3 coupled with pilot selector No. 2 is used to control the sign for the odd numbered rows of the reduced correlation matrix while pilot selector No. 4 coupled with pilot selector No. 2 is used to control the sign for the even numbered rows. For each successive multiplication the sign control for the previous multiplication must be dropped out. This is done by impulsing the digit selector from the program exit. The digit selector emits a timed impulse for 9, 8, 7, 6, 5, 4, 3, 2, and 1. No impulse is emitted for zero. The 9 hub of the digit selector is wired to drop out hubs of pilot selectors 2, 3, and 4. In this way the pilot selectors are reset before the 5 impulse is emitted from the storage units.

Pilot selectors are double pole double throw relays. Each selector has two switching elements which work together mechanically but are electrically independent. The diagram below shows schematically a pilot selector. The two sections (a) and (b) constitute the switching features of one pilot selector. If the pilot selector is not energized the C or common hub of both section (a) and (b) are connected to the N or normal hub. When the selector is energized the armature swings to the dotted position shown in the diagram and both sections are now connected to the T or transferred hub.

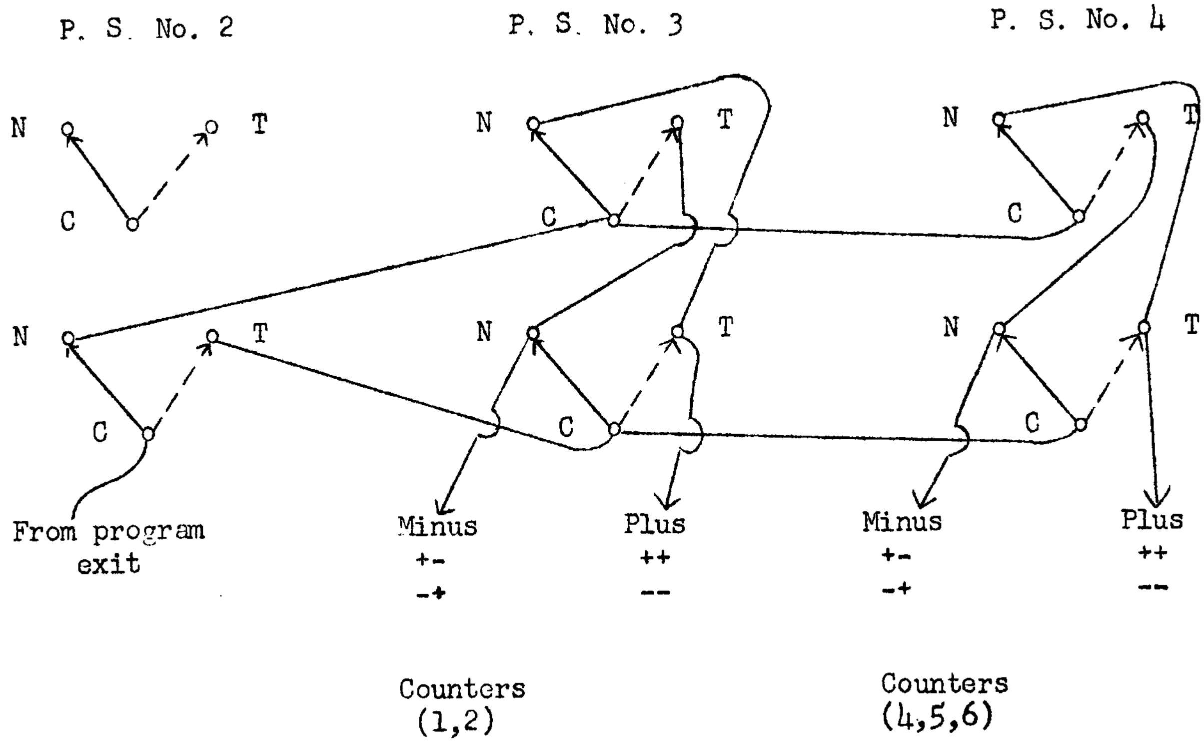
(a)



(b)



This makes it possible in this problem for us to control the sign of the two products developed simultaneously with only three selectors. Figure 4 shows the connections through the pilot selectors so as to add or subtract into the counters



The development of one product will be followed through the pilot selectors. Pilot selectors No. 2, No. 3, and No 4 are returned to normal position at the beginning of the sign control program. If u_1 is negative a 5 punch in the left hand position of the storage unit 6L causes pilot selector No. 2 to transfer to the position shown by the dotted line. h_{11}^2 must be positive so pilot selector No. 3 is not energized. The impulse from the program exit hub is connected to the common hub of P. S. No. 2 and exits from the T hub and is then wired to the C hub of P. S. No. 3 coming out in this case of the N hub of P. S. No. 3 and then to counters 1, 2 where it is wired to subtract.

Program Step 2:

The multiplier u_1 is read out of storage unit 1R while the multiplicands h_{11}^2 and r_{21} are read out of storage units 2L and 3R respectively. The machine is impulsed to multiply. The product $u_1h_{11}^2$ is developed and added or subtracted into counters 1 and 2 depending upon the setting of pilot selectors No. 2 and No. 3. The product u_1r_{21} is developed and added or subtracted into counters 4, 5, and 6 depending upon the setting of pilot selectors No. 2 and No. 4.

Program Step 3:

This is the second sign control program. Pilot selectors No. 2, No. 3, and No. 4 are reset. The left hand position of storage unit 7L is read out to pilot selector No.2 while the remaining part of storage unit 7L is read out and read into storage unit 1R. The left hand position of storage unit 2R is read out to pilot selector No. 3 while the left hand position of storage unit 4L is read out to pilot selector No. 4.

Program Step 4:

The multiplier u_2 is read out of storage unit 1R while the multiplicands r_{12} and h^2_{22} are read out of storage units 2R and 4L respectively. The machine is impulsed to multiply. The product u_2r_{12} is developed and added or subtracted into counters 1 and 2 depending upon the setting of pilot selectors No. 2 and No. 3. The product $u_2h^2_{22}$ is developed and added or subtracted into counters 4, 5, and 6 depending upon the setting of pilot selectors No. 2 and No. 4.

Program Step 5:

This is the third sign control program. Pilot selectors No. 2, No. 3 and No. 4 are reset. The left hand position of storage unit 7R is read out to pilot selector No. 2 while the remaining part of storage unit 7R is read out to storage unit 1R. The left hand position of storage unit 3L is read out to pilot selector No. 3 while the left hand position of storage unit 4R is read out to pilot selector No. 4

Program Step 6:

The multiplier u_3 is read out of storage unit 1R while the multiplicands r_{13} and r_{23} are read out of storage units 3L and 4R respectively. The machine is impulsed to multiply. The product u_3r_{13} is developed and added or subtracted into counters 1 and 2 depending upon the setting of pilot selectors No. 2 and No. 3. The product u_3r_{23} is developed and added or subtracted into counters 4, 5, and 6, depending upon the setting of pilot selectors No. 2 and No. 4.

Program Step 7:

The machine is impulsed to read the next card. The first matrix card skips out and the second multiplier card is read into the machine. The remaining four multiplier cards with their corresponding matrix cards pass through the machine in the same manner. When the last matrix card, No. 5, has been computed, the values v_1 and v_2 have been accumulated in counters 1 and 2, and counters 4, 5, and 6. Blank trailer cards were placed after every fifth matrix card This trailer card has an 11 or X punch in column 79. This X punch is sensed by the reading brushes and causes the machine to skip the first sever program steps.

Program Step 8:

The code number is read out of counter 3 and the counter is reset to zero. This number is read into storage unit 6R. Storage unit 6R is impulsed to read out and to punch into columns 1-4 of the trailer card. The code number for the first trailer card is 0101.

Program Step 9:

The product v_1 is read out of counters 1 and 2 and read into 6L and 6R. Counters 1 and 2 are reset to zero. The ninth hub from the right of storage units 6L and 6R combined is wired through a balance test hub to determine the sign of the product. If the sign is negative an 11 of X punch is wired to punch in the column containing the first digit. Storage units 6L and 6R are read out and impulsed to punch into columns 61-68.

Program Step 10:

The product v_2 is read out of counters 4, 5, and 6 and read into storage units 7L and 7R. Counters 4, 5, and 6 are reset to zero. The ninth hub from the right of storage units 7L and 7R combined is read through the balance test hub to determine the sign. An 11 or X punch is punched over the left hand digit of the result to indicate a negative value. Storage units 7L and 7R are read out and impulsed to punch into columns 69-76.

Program Step 11:

The machine is impulsed to read the next card.

This completes the computation of the first two elements of vector V. The remaining elements are computed in a similar manner. The actual running time for a single iteration depends on the number of matrix cards involved. It requires approximately eight seconds to develop one set of products such as $u_1h_{11}^2$ and u_1r_{21}. For the example given in Figure 1 there are forty matrix cards. The time required would be 320 seconds plus approximately two seconds for the punching of each trailer card or an additional 16 seconds making a total time of 336 seconds. This would be 5 minutes and 36 seconds.

After the eight trailer cards have been punched, these can be separated out by sorting on column 79. The trailer cards will all fall into the 11 pocket. These cards are listed and summed on the tabulator. The sum must check rather closely with the computed summational check.

The largest entry is used as a divisor for all entries in the V column vector. This constitutes the second approximation of the U vector. A new set of multiplier cards is made up and the multiplication of R x U is carried out again.

Disadvantages of the Method:

Because of inadequate storage capacity, a considerable number of cards must be used to repeatedly read in the multiplier values. In the sample problem forty multiplier cards were required. In a thirty variable problem 160 multiplier cards would be needed. This number of cards would have to be used for each iteration. In many instances a large number of iterations would be required. The number of cards used might well become extremely large.

Another disadvantage is the extremely slow convergence rate of the iterative solution when two or more roots of the reduced correlation matrix are approximately equal. In the next section of this report a discussion of an alternative solution is given which tends to speed up the rate of convergence. The same basic plan for using the 602 A can be utilized for this alternative solution.

3. An iterative method for the simultaneous extraction of two or more roots:

A. History of the Method.

Because of the disadvantages of Hotelling's iterative method as outlined in Section I, a great deal of time and effort have gone into investigations of new techniques which were designed to reduce substantially the computational time involved. Mr. R. H. Morris of Eastman

Kodak Company became interested in the problem about five years ago and attempted to find a new method which would satisfy this intent. To describe his approach it is necessary to indicate the type of data with which he was working.

The manufacture and control of color film is a very complicated process; it is necessary to take observations on a large number (say 40 or 50) of highly correlated variables. For theoretical and practical reasons it is thought that several linear functions of these variables will suffice for control purposes and also give some indication of the underlying chemical reactions which occur in the film. The theoretical foundation of this approach lies in the fact that at least three dyes are known to exist in the film and theory postulates that these three dyes have an additive effect (i.e., $Y = ax_1 + bx_2 + cx_3$). Under this hypothesis the model becomes precisely that of principal component theory.

In colorimetry the most accurate quantitative data is obtained by a spectrophotometric curve of the color patch to be measured. This consists in measuring the transmittance (or reflectance) of the material across the visible spectrum from 400-700 millimicrons at 10 millimicron intervals. The procedure then involves taking a large number of patches, corresponding to N or sample size, and computing the 31 x 31 covariance matrix taking as the 31 variables the wavelengths at 10 mμ intervals from 400-700 mμ. Covariance matrices are used here since the variables are all in the same units; thus there is no need to use correlation matrices. Still there is no essential difference in the method between the two approaches. If in the analysis more than three vectors were obtained, this may indicate the non-additivity of the process or the formation of some new dyes. This finding is then reported to the physicists and chemists who explore it further.

Perhaps the most effective use of principal components is in the quality control field where an almost impossible number of control charts would otherwise be necessary. Controls are kept on the values of the three or four linear functions themselves employing Hotelling's methods as outlined in Techniques of Statistical Analysis. Moreover, investigation and indeed determination of new films are facilitated by this technique which, as shown above, saves a great deal of labor and still gives results sufficiently accurate for most applications.

Hotelling's method was therefore employed on the covariance matrices of these variables. In the majority of cases it was found that three or four vectors were sufficient to remove 95-99 o/o of the trace of the covariance matrix and that, although convergence of the first vector was fairly rapid, the remaining ones converged comparatively slowly.

Because of the great amount of computation necessary, Morris then devised a method of obtaining simultaneously the vectors corresponding to the k largest roots of a non-negative symmetric matrix A of order n x n .

B. Description of the Method.

Take an arbitrary k x n semi-orthogonal matrix X_0 [i.e., $X_0X_0' = I(k)$, X_0 consists of the first k rows of an orthogonal matrix of order n x n] Form $X_0A = Y_0$ (k x n). The rows of Y_0 will be "closer" to the first k characteristic vectors of A than those of X_0 in the sense that the rows of Y_0, considered as vectors, are nearer to the space generated by the k characteristic vectors. This follows from the section on powering the matrix. In fact, if A is actually of rank k, the rows of Y_0 will lie in the space of the k characteristic vectors. However, the rows of Y_0 are no longer orthonormal so that if Y_0 is left unmodified and repeatedly pre-multiplied by A, all k rows will converge to the first characteristic vector of A. This can be altered to give convergence to the first k vectors of A by making the rows of Y_0 mutually orthogonal and of unit length. Thus, Y_0 is modified to obtain X_1 such that $X_1X_1' = I(k)$. Y_0 must be modified so that the rows of Y_0 and X_1 generate the same space and this requirement is equivalent to $X_1 = T_0(k \times k)Y_0$ where T_0 is non-singular.

Then the iterative scheme is

$$Y_i = X_iA$$

$$X_{i+1} = T_iY_i$$

and at each step,

$$X_iX_i' = I(k) \qquad |T_i| \neq 0$$

The iteration is terminated when T_i becomes a diagonal matrix.

The problem remaining is to define T. This can be done in an infinity of ways, for if any matrix M is found that will orthonormalize the rows of Y, then NM where N is any k x k orthogonal matrix will also orthonormalize the rows for

$$MYY'M' = I(k)$$

so

$$NMYY'M'N' = NN' = I(k) \ .$$

Now there are two criteria for defining T:

(1) Convenience in computation or programming,

(2) Speed of convergence.

Method I.

The most convenient method is the Gram-Schmidt or Square root process which is as follows:

Consider

$$\begin{matrix} YY'(k \times k) & I(k) \\ K(k \times k) & L(k \times k) \end{matrix}$$

where K and L are upper and lower triangular matrices ($t_{ij} = 0$ for $i > j$ for upper triangular matrix and $t_{ij} = 0$ for $i < j$ for lower). $K'K = YY'$ and $K'L = I(k)$. Then L may be taken as T, for if $X = LY$ $XX' = LYY'L' = LK'KL' = (LK')^2 = I(k)I(k) = I(k)$. This is the standard square root method as outlined in P. S. Dwyer's papers and others: it takes the first row of X to be the first row of Y normed to unit length; the second row of X is a linear combination of the first two rows of Y such that the first two rows of X are orthogonal and of unit length, etc.

Method II.

The fastest convergence as shown by empirical evidence seems to be obtained by defining T as the appropriately normalized k x k matrix of characteristic vectors of YY'(k x k).

Let V be the matrix having as its rows the characteristic vectors of YY', each normed to unit length and let Λ be the diagonal matrix whose elements are the corresponding roots. Then $V(YY') = \Lambda V$ and $VV' = I(k)$. Take $T = \Lambda^{-\frac{1}{2}} V$. Thus if $X = \Lambda^{-\frac{1}{2}} VY$,

$$XX' = \Lambda^{-\frac{1}{2}} V(YY')V' \Lambda^{-\frac{1}{2}} = \Lambda^{-\frac{1}{2}} \Lambda VV' \Lambda^{-\frac{1}{2}} = I(k) \quad .$$

This method, though involving less iteration, will be much more difficult to program than Method I on computers such as the IBM 602A, 604 and CPC. This results from a lack of storage capacity on these machines together with the difficulty in extracting the roots and vectors of a k x k matrix, even one as small as 4 x 4.

Morris uses Method I routinely employing the CPC.

C. Discussion of the Method.

The chief advantages of this method are first, elements of A are not altered by computation of residuals, thus helping control bounding error; and second, speed of convergence is much increased, particularly if the roots are close together. The last statement applies especially to Method II.

However, it seems that using Method II, the off diagonal elements of YY' become so small with respect to the diagonal elements, that it is rather difficult to extract the characteristic vectors of YY', at least on a desk machine. This can be avoided by, at this point, assuming that the vector associated with the largest root is sufficiently stable, as it should be at this stage, and then using Method I to get the other orthogonal rows of X.

The initial starting matrix X_0 might be a little less arbitrary. If all the correlations are positive, it would be advisable to use $1/\sqrt{N}$ as the elements of the first row of X_0. Supposing the column sums of A to be employed as the first row of X_0, there would be some difficulty in finding the other rows of X_0 such that they are orthogonal to the first row. Therefore this method is not recommended.

On the other hand, the choice of the lower k - 1 rows seems to be indeterminate. Determining the sequence of signs in the other rows by using the centroid method of factoring would appear of little value because of the great discrepancy between characteristic vectors and centroid factors following the first.

At Kodak, Fisher and Yates' orthogonal polynomials are used as X_0 when the correlations are all positive.

A mathematical description of Method I, employing a 2 x n matrix as X_0, gives an insight into the advantages and disadvantages of this procedure. Using the notation of the section on powering the matrix, consider the first row of X_0 [or the first column of X_0'] as the transpose of a linear combination of the n true characteristic vectors, i.e., $(a_1V_1 + .. + a_nV_n)'$, and in the same manner take the second row of X_0 to be $(b_1V_1 + .. + b_nV_n)'$. Then

$$A(\text{1st column of } X_0') = a_1AV_1 + .. + a_nAV_n = a_1\lambda_1V_1 + \dots + a_n\lambda_nV_n = Y_1$$

$$A(\text{2nd column of } X_0') = b_1AV_1 + \dots + b_nAV_n = b_1\lambda_1V_1 + \dots + b_n\lambda_nV_n = Y_2$$

If the first and second rows of X_0 had been postmultiplied by A, the result would be the transpose of the above, and the calculations would come out the same.

Now, proceeding as in Method I,

$$\begin{bmatrix} a_1^2\lambda_1^2 + \dots + a_n^2\lambda_n^2 & a_1b_1\lambda_1^2 + \dots + a_nb_n\lambda_n^2 \\ a_1b_1\lambda_1^2 + \dots + a_nb_n\lambda_n^2 & b_1^2\lambda_1^2 + \dots + b_n^2\lambda_n^2 \end{bmatrix} = \begin{bmatrix} (1) & (2) \\ (2) & (3) \end{bmatrix}$$

$$K = \begin{bmatrix} \sqrt{(1)} & \dfrac{(2)}{\sqrt{(1)}} \\ 0 & \sqrt{(3) - \dfrac{(2)^2}{(1)}} \end{bmatrix}$$

$$L = \begin{bmatrix} \dfrac{1}{\sqrt{(1)}} & 0 \\ \\ -\dfrac{(2)}{(1)\cdot\sqrt{(3) - \dfrac{(2)^2}{(1)}}} & \dfrac{}{\sqrt{(3) - \dfrac{(2)^2}{(1)}}} \end{bmatrix}$$

Consider the second row of L operating on Y_1 and Y_2, or

$$Y_2' = \left[(3) - \frac{(2)^2}{(1)} \right]^{-\frac{1}{2}} \left\{ - \frac{a_1 b_1 \lambda_1^2 + \ldots + a_n b_n \lambda_n^2}{a_1^2 \lambda_1^2 + \ldots + a_n^2 \lambda_n^2} \; Y_1 + Y_2 \right\} =$$

$$\left[(3) - \frac{(2)^2}{(1)} \right]^{-\frac{1}{2}} \left\{ - \frac{\dfrac{b_1}{a_1} + \dfrac{a_2 b_2}{a_1^2} \left(\dfrac{\lambda_2}{\lambda_1}\right)^2 + \ldots + \dfrac{a_n b_n}{a_1^2} \left(\dfrac{\lambda_n}{\lambda_1}\right)^2}{1 + \dfrac{a_2^2}{a_1^2} \left(\dfrac{\lambda_2}{\lambda_1}\right)^2 + \ldots + \dfrac{a_n^2}{a_1^2} \left(\dfrac{\lambda_n}{\lambda_1}\right)^2} \; Y_1 + Y_2 \right\} \approx$$

$$\left\{ - \frac{b_1}{a_1} (a_1 \lambda_1 V_1 + a_2 \lambda_2 V_2 + \ldots + a_n \lambda_n V_n) + b_1 \lambda_1 V_1 + \ldots + b_n \lambda_n V_n \right\} \left[(3) - \frac{(2)^2}{(1)} \right]^{-\frac{1}{2}}$$

$$= \left[(3) - \frac{(2)^2}{(1)} \right]^{-\frac{1}{2}} \left\{ \left(b_2 - \frac{b_1 a_2}{a_1}\right) \lambda_2 V_2 + \ldots + \left(b_n - \frac{b_1 a_n}{a_1}\right) \lambda_n V_n \right\}$$

Therefore, to a certain order of approximation V_1 has been removed; then on iteration the result will lie in a space orthogonal to V_1.

From the above discussion for Method I, it is evident that the number of iterations to obtain V_1 to the desired accuracy is not reduced by this method, since the first row is left unchanged by the transformation and iterated. It is also apparent that the convergence of the second vector is still a function of the third root and will not in general be changed (i.e., same number of iterations) from that of the usual method; however, residual matrices need not be computed and this is where the saving arises.

The results may be generalized to the case where X_0 contains more than two rows. For example, if X_0 contains three rows, then the speed of convergence to the second vector should be considerably quickened, while for the third the number of iterations would remain approximately the same, but the two residual matrices need not be computed. Thus it seems advantageous when using Method I to set $k = 4$ at least and proceed as outlined.

On the other hand, if there is a reliable estimate of the rank of the matrix based on theory or experience, Method II would give the quickest convergence, but as mentioned before, if k is rather large the work involved would still be appreciable. In this case, it amounts to using principal components on a $k \times k$ rather than an $n \times n$ matrix.

Since in psychological data k will usually be unknown, Method I is recommended for this reason and for the greater convenience in programming.

There are as yet no published accounts of this method, Morris having given it only in discussion form at an IBM seminar at Endicott.

D. Sample Problems.

A small fictitious example constructed by Morris will illustrate the possible saving. Method II was employed on the following matrix:

$$\begin{bmatrix} 3.150108 & .050108 & .950108 & .850108 \\ & 4.350108 & -.349892 & .950108 \\ & & 4.350108 & .050108 \\ & & & 3.150108 \end{bmatrix}$$

$\lambda_1 = 5.0004321$

$\lambda_2 = 5.0000000$

$\lambda_3 = 3$

$\lambda_4 = 2$

Starting with $X_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, the multiple extraction procedure took 23 iterations while Hotelling's iterative method would take better than 20,000 iterations to get five place accuracy on the first vector. Though obviously a situation such as this will not usually occur in practice, this may give some indication of relative efficiency.

In order to give a more practical example a small problem has been undertaken here. It consists of twelve body measurements; the data is fallible and has been factored previously by the centroid method in which four factors were extracted. The diagonal elements of the reduced correlation matrix were taken as the communalities from the previous study. As a starting point an orthogonal matrix of order 2 x 12 was set up; the elements of the first row being all $1/\sqrt{12}$ and in the second row, the first half $1/\sqrt{12}$ and the second half $-1/\sqrt{12}$.

To compare Hotelling's method with the multiple extraction procedure, the first vector was determined by Hotelling's method. The starting vector was selected proportional to the column sums of the reduced correlation matrix. After twelve iterations two successive u's were the same to six decimals. Using the multiple extraction method the first vector was also found accurate to six decimals after twelve trials while the second axis had not yet become stable. A number of additional iterations will be necessary to gain convergence. This is due to the relative closeness of the second and third roots; the situation would no doubt be improved by using a 3 x 12 matrix as X_1 rather than a 2 x 12.

4. <u>Diagonal Estimates for the Reduced Correlation Matrix</u>.

To date no solution to this problem has been obtained, though some ideas have been advanced and are at present being investigated. Guttman in a recent paper discussed the different estimates which could be used. He recommends that the multiple correlation between each variable and the other n-1 variables be used as the estimate. The computational labor involved in a large study would make this method seem impractical.

W. G. Howe has proposed that the "best" estimates of these communalities are those obtained by factorizing the unreduced correlation matrix (with one's in the diagonals,R_1),preferably using characteristic vectors. Then, after these vectors, $V(k \times n)$, have been so normalized that $VV' = \lambda$, the diagonal elements of $V'V$ are selected and inserted in the diagonals of R_1. The new matrix R is iterated until a new set of characteristic vectors, say V_1, is obtained. However, the computation here will be much less, since V itself may be used as the initial approximation and should converge rapidly to V_1. This procedure is then continued until the diagonal elements become stable; to wit $V_i'V_i = V_{i+1}'V_{i+1}$ or $V_i = V_{i+1}$.

The scheme as outlined above has a bad drawback in that the rank of R (actually the rank of the population reduced correlation matrix) is unknown. In the above it has been assumed that $V'V = R$, or that k vectors remove enough of the variability. But since this rank is unknown, it is very difficult to tell at what stage to stop factoring. Investigation is being conducted along these lines.

However, at present the highest entry in each column is the estimate of the communality that has been used here.

R

h^2_{11}	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}	r_{18}	r_{19}	$r_{1,10}$	$r_{1,11}$	$r_{1,12}$	$r_{1,13}$	$r_{1,14}$	$r_{1,15}$
Matrix Card #1			Matrix Card #2			Matrix Card #3			Matrix Card #4			Matrix Card #5		
r_{21}	h^2_{22}	r_{23}	r_{24}	r_{25}	r_{26}	r_{27}	r_{28}	r_{29}	$r_{2,10}$	$r_{2,11}$	$r_{2,12}$	$r_{2,13}$	$r_{2,14}$	$r_{2,15}$
r_{31}	r_{32}	h^2_{33}	r_{34}	r_{35}	r_{36}	r_{37}	r_{38}	r_{39}	$r_{3,10}$	$r_{3,11}$	$r_{3,12}$	$r_{3,13}$	$r_{3,14}$	$r_{3,15}$
Matrix Card #6			Matrix Card #7			Matrix Card #8			Matrix Card #9			Matrix Card #10		
r_{41}	r_{42}	r_{43}	h^2_{44}	r_{45}	r_{46}	r_{47}	r_{48}	r_{49}	$r_{4,10}$	$r_{4,11}$	$r_{4,12}$	$r_{4,13}$	$r_{4,14}$	$r_{4,15}$
r_{51}	r_{52}	r_{53}	r_{54}	h^2_{55}	r_{56}	r_{57}	r_{58}	r_{59}	$r_{5,10}$	$r_{5,11}$	$r_{5,12}$	$r_{5,13}$	$r_{5,14}$	$r_{5,15}$
Matrix Card #11			Matrix Card #12			Matrix Card #13			Matrix Card #14			Matrix Card #15		
r_{61}	r_{62}	r_{63}	r_{64}	r_{65}	h^2_{66}	r_{67}	r_{68}	r_{69}	$r_{6,10}$	$r_{6,11}$	$r_{6,12}$	$r_{6,13}$	$r_{6,14}$	$r_{6,15}$
r_{71}	r_{72}	r_{73}	r_{74}	r_{75}	r_{76}	h^2_{77}	r_{78}	r_{79}	$r_{7,10}$	$r_{7,11}$	$r_{7,12}$	$r_{7,13}$	$r_{7,14}$	$r_{7,15}$
Matrix Card #16			Matrix Card #17			Matrix Card #18			Matrix Card #19			Matrix Card #20		
r_{81}	r_{82}	r_{83}	r_{84}	r_{85}	r_{86}	r_{87}	h^2_{88}	r_{89}	$r_{8,10}$	$r_{8,11}$	$r_{8,12}$	$r_{8,13}$	$r_{8,14}$	$r_{8,15}$
r_{91}	r_{92}	r_{93}	r_{94}	r_{95}	r_{96}	r_{97}	r_{98}	h^2_{99}	$r_{9,10}$	$r_{9,11}$	$r_{9,12}$	$r_{9,13}$	$r_{9,14}$	$r_{9,15}$
Matrix Card #21			Matrix Card #22			Matrix Card #23			Matrix Card #24			Matrix Card #25		
$r_{10,1}$	$r_{10,2}$	$r_{10,3}$	$r_{10,4}$	$r_{10,5}$	$r_{10,6}$	$r_{10,7}$	$r_{10,8}$	$r_{10,9}$	$h^2_{10,10}$	$r_{10,11}$	$r_{10,12}$	$r_{10,13}$	$r_{10,14}$	$r_{10,15}$
$r_{11,1}$	$r_{11,2}$	$r_{11,3}$	$r_{11,4}$	$r_{11,5}$	$r_{11,6}$	$r_{11,7}$	$r_{11,8}$	$r_{11,9}$	$r_{11,10}$	$h^2_{11,11}$	$r_{11,12}$	$r_{11,13}$	$r_{11,14}$	$r_{11,15}$
Matrix Card #26			Matrix Card #27			Matrix Card #28			Matrix Card #29			Matrix Card #30		
$r_{12,1}$	$r_{12,2}$	$r_{12,3}$	$r_{12,4}$	$r_{12,5}$	$r_{12,6}$	$r_{12,7}$	$r_{12,8}$	$r_{12,9}$	$r_{12,10}$	$r_{12,11}$	$h^2_{12,12}$	$r_{12,13}$	$r_{12,14}$	$r_{12,15}$
$r_{13,1}$	$r_{13,2}$	$r_{13,3}$	$r_{13,4}$	$r_{13,5}$	$r_{13,6}$	$r_{13,7}$	$r_{13,8}$	$r_{13,9}$	$r_{13,10}$	$r_{13,11}$	$r_{13,12}$	$h^2_{13,13}$	$r_{13,14}$	$r_{13,15}$
Matrix Card #31			Matrix Card #32			Matrix Card #33			Matrix Card #34			Matrix Card #35		
$r_{14,1}$	$r_{14,2}$	$r_{14,3}$	$r_{14,4}$	$r_{14,5}$	$r_{14,6}$	$r_{14,7}$	$r_{14,8}$	$r_{14,9}$	$r_{14,10}$	$r_{14,11}$	$r_{14,12}$	$r_{14,13}$	$h^2_{14,14}$	$r_{14,15}$
$r_{15,1}$	$r_{15,2}$	$r_{15,3}$	$r_{15,4}$	$r_{15,5}$	$r_{15,6}$	$r_{15,7}$	$r_{15,8}$	$r_{15,9}$	$r_{15,10}$	$r_{15,11}$	$r_{15,12}$	$r_{15,13}$	$r_{15,14}$	$h^2_{15,15}$
Matrix Card #36			Matrix Card #37			Matrix Card #38			Matrix Card #39			Matrix Card #40		
Sr_{j1}	Sr_{j2}	Sr_{j3}	Sr_{j4}	Sr_{j5}	Sr_{j6}	Sr_{j7}	Sr_{j8}	Sr_{j9}	Sr_{j10}	Sr_{j11}	Sr_{j12}	Sr_{j13}	Sr_{j14}	Sr_{j15}

• U = V

U
u_1
u_2
u_3
u_4
u_5
u_6
u_7
u_8
u_9
u_{10}
u_{11}
u_{12}
u_{13}
u_{14}
u_{15}

V
v_1
v_2
v_3
v_4
v_5
v_6
v_7
v_8
v_9
v_{10}
v_{11}
v_{12}
v_{13}
v_{14}
v_{15}
Check Sum

Figure 1

APPLICATION ______ PROBLEM ______

IBM CALCULATING PUNCH TYPE 602 A PLANNING CHART

PROGRAM SUPPRESS	PROGRAM STEP	OPERATION	STORAGE UNIT 1L	STORAGE UNIT DIVR.–MULT. 1R	COUNTER DIVIDEND 1	COUNTER DIVIDEND 2	COUNTER DIVIDEND 3	COUNTER 4	COUNTER 5	COUNTER 6	STORAGE UNITS 2L	2R	3L	3R	4L	4R	PUNCH UNITS* 6L	6R	7L	7R	PROGRAM STEP	
		READ					R.I. Code (1 - 4)										R.I. u_1 (41 - 46)		R.I. u_2 (47 - 52)	R.I. u_3 (53 - 58)		X-78 Detail Card Skip O.
		CYCLE									R.I. h_{11}^2 (5 - 10)	R.I. r_{12} (11 - 16)	R.I. r_{13} (17 - 22)	R.I. r_{21} (23 - 28)	R.I. h_{22}^2 (29-34)	R.I. r_{23} (35-40)						X-71 Detail Card Skip O.
	1	Sign Control		R.I. from 6L							R.O. to P.S. #3			R.O. to P.S. #4			R.O. to 1R and P.S. #2				1	
	2	$u_1 \times h_{11}^2$ $u_1 \times r_{21}$		R.O. Mult.	R.I. $\pm(u_1 \times h_{11}^2)$				R.I. $\pm(u_1 \times r_{21})$		R.O. to (1, 2)			R.O. to (4, 5, 6)							2	
	3	Sign Control		R.I. from 7L								R.O. to P.S. #3			R.O. to P.S. #4				R.O. to 1R and P.S. #2		3	
	4	$u_2 \times r_{12}$ $u_2 \times h_{22}^2$		R.O. Mult.	R.I. $\pm(u_2 \times r_{12})$				R.I. $\pm(u_2 \times h_{22}^2)$			R.O. to (1, 2)			R.O. to (4, 5, 6)						4	
	5	Sign Control		R.I. from 7R									R.O. to P.S. #3			R.O. to P.S. #4				R.O. to 1R and P.S. #2	5	
	6	$u_3 \times r_{13}$ $u_3 \times r_{23}$		R.O. Mult.	R.I. $\pm(u_3 \times r_{13})$				R.I. $\pm(u_3 \times r_{23})$				R.O. to (1, 2)			R.O. to (4, 5, 6)					6	
	7	Impulse Read																			7	
	8	Punch Code No.					R.O. and Reset											R.I. (3) Punch (1 - 4)			8	X-79 Trailer Card Only
	9	Punch $\sum r_{1j}u_{j1}$			R.O. and Reset												R.I. (1, 2) Punch (61 - 68)				9	"
	10	Punch $\sum r_{2j}u_{j1}$							R.O. and Reset										R.I. (4, 5, 6) Punch (69 - 76)		10	"
	11	Impulse Read																			11	
	12																				12	

*UNITS POSITION MUST BE WIRED TO PUNCH

FORM 22-9323-1

Figure 2

INTERNATIONAL BUSINESS MACHINES CORPORATION

IBM CALCULATING PUNCH
TYPE 602 A CONTROL PANEL

FORM 22 9323 1 PRINTED IN U.S.A.

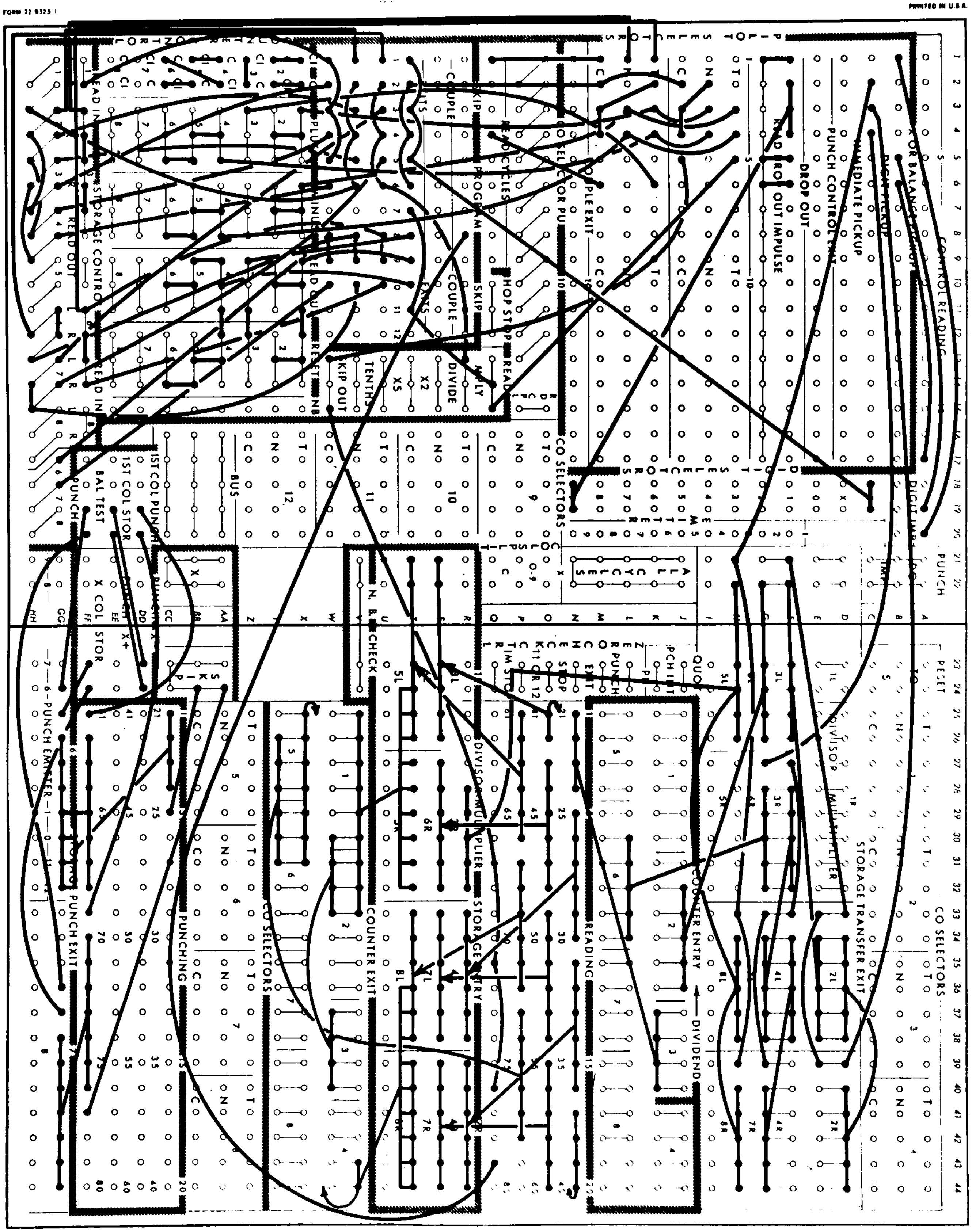


Figure 3