CHAPTER XIII

THE FACTOR ANALYSIS OF THE TEST DATA

> Mental testing fails to provide scientific measurement of human abilities.......because its results are so far incapable of fulfilling the essential conditions of scientific quantification.
>
> - Thomas, L. G.[1]

One of the essential conditions of measurement that Thomas must have in his mind is that of univocalness of the meaning of scores.

Guilford[2] says,

> Factor theory has highlighted a very serious fault in psychological tests. This is the fact that any test measures more than one common factor to a substantial degree yields scores that are psychologically ambiguous and very difficult to interpret. What is worse, almost all tests have a complexity greater than one, that is, they measure more than one common factor.

The same author[3] adds,

> The fundamental variables or dimensions of human ability and of human personality in general are

---

1 Thomas, L. G., Mental tests as instruments of Science, Psychol. Monogr., 1942, 54, No. 245.

2 Guilford, J.P., "Psychometric Methods", McGraw-Hill Book Company, Inc., New York, 1954, p. 356.

3 Ibid., p. 470.

still well within the unexplored
territory reserved for psychologists.

To meet this situation - to search for the unitary traits of personality - a statistical approach such as factor analysis is necessary.

The shortcomings of commonly used single score tests have been revealed by the statistical procedure known as factor analysis. This procedure was first developed and applied to mental ability tests by Spearman and his co-workers in England. In the United States, pioneer work in this field was conducted by Thurstone.

Following the work of Spearman and Thurstone, many other investigators have used factor analysis to study the nature of human abilities. Some of these studies have been used merely to gain a greater understanding of the organisation and components of such abilities. Others have been used as the foundation for the construction of multifactor tests. Through factor analysis, psychologists have contributed much to the understanding of tests and of human behaviour.

Application of factor-analysis techniques to results of test performance has provided a means of making a direct attack upon the problem of test validity. Authors of tests are becoming increasingly concerned about the extent to which various parts of their tests, as well as the tests as a whole, isolate and measure, as such, relatively independent factors.

Factor-analysis techniques are now widely applied to varied kinds of instruments, such as tests of motor skill, intelligence tests, special aptitude tests, and personality inventories. As a result of their use, authors of tests are able to provide those, who use them with more exact information concerning their component elements than formerly.

Factor analysis has also been applied in such varying fields as physical measurement, political science, finance, and medicine. The method is essentially a statistical tool.

GENERAL NATURE OF FACTOR ANALYSIS

Factor analysis as applied to human behaviour is somewhat analogous to quantitative analysis in chemistry, where a compound is analysed to discover the nature and amounts of its basic elements.

"In factor analysis", as B. Fruchter[1] says,

> A series of test scores or other measures are inter-correlated to determine the number of dimensions the test space occupies, and to identify these dimensions in terms of traits or other general concepts.

Again in the words of Holzinger and Harman[2],

> Factor analysis is a branch of statistical theory concerned with the

---

1 Fruchter, Benjamin, "Introduction to Factor Analysis", D. Van Nostrand Company, Inc., Princeton, New Jersey, 1954, p.2.

2 Holzinger, K.J., & Harman, H.H., "Factor Analysis", University of Chicago Press, Chicago, 1941, p.3.

resolution of a set of descriptive variables in terms of a smaller number of categories or factors. This resolution is accomplished by the analysis of the inter-correlations of the variables.

DIFFERENT METHODS OF FACTOR-ANALYSIS

Since Spearman proposed his criterion of the tetrad difference, a number of procedures for factor analysis have been proposed. All of them start with the same kind of data - a correlation matrix.

The procedures for extracting factors that are most commonly used are listed below.

(i) Method of principal components of Hotelling.

(ii) Method of principal axes of Kelley.

(iii) Method of summation of Burt.

(iv) The centroid method of Thurstone.

The first two have much in common, and also the second two.

In addition to the process of extracting factors, Thurstone provides rotation methods to arrive at meaningful factors. Holzinger has designed methods of solving for factors in which he requires that every test yields some relationship to a 'g' factor and to one, but only one, group factor. In the United Kingdom, the investigators generally follow Burt's method of summation while in the U.S.A., Thurstone's centroid method is almost invariably used.

'Factor-analysis' is still quite a new technique to the research workers in this country. Uptil now quite a few investigators have applied this technique to their investigations. And naturally, a fewer amongst them might have an expert knowledge of the whole technique.

The present investigator had no knowledge of the technique at all at the time of taking up this project. Following the suggestion of his guiding professor, he decided to apply Thurstone's centroid method to the present test data and acquired a working knowledge of the technique. Even then it should be noted that his knowledge of the technique is mainly pertaining to the use of centroid method only.

As quoted by Vernon,[1] Thomson has pointed out that when tests are given with the object of predicting educational or vocational achievement, the extraction of factors from the tests may constitute quite an unnecessary intermediate step. A much more efficient method is to correlate the tests directly with measures of achievement and then to use multiple correlation for obtaining the most accurate predictions.

The Wherry-Doolittle method for selecting tests in a battery has been applied to the present tests and the use of

1 Vernon, P. E., "The Measurement of Abilities", University of London Press Ltd., London, 1956, Footnote on page 145.

multiple correlation for obtaining predictions is also made.

Thomson's view is, of course, perfectly true from the statistical standpoint. But following the recent trends in mental testing, the present investigator feels that it is necessary to apply this modern technique to the present investigation to make it more thorough and more useful.

THURSTONE'S CENTROID METHOD OF FACTORING

The Thurstone theory and methods have been developed on the basis of matrix algebra.

The term "centroid" as used in factor analysis is closely allied with its mechanical concept. In mechanics, the centroid is a point in a mass where the centre of gravity is located. In factor analysis, the centroid of the end points of the test vectors might be considered the location of the centre of gravity of equal weights at the points. A centroid is, then, a centre of gravity. Statistically regarded, it is a mean.

As is stated earlier, the purpose of factoring a correlation matrix is to account for the inter-correlations with fewer factors than there are tests. This factoring should be done so as to minimise the residuals after each factor has been determined. The main centroid axis is regarded as an approximation to the major principal axis of the factor configuration. This main centroid axis is so placed that it has zero projections on all the remaining coordinate axes. This fact leads to the theorem that "The sum of the coefficients in the cprrelation matrix is equal to the square of the sum of the first centroid factor loadings,"[1] permitting factoring through simple summational procedure after appropriate reflections. By "reflecting" is meant that each test vector retains its same length but it extends in the opposite direction. The general policy is to reflect one test vector at a time and note the results; then reflect a second one, and so on.

The extraction of each factor loading reduces the residuals in the correlational matrix. The factoring process is ordinarily stopped when the standard deviation of the residuals is less than the standard error of a zero correlation.

REPORTING THE FACTORIAL STUDY OF THE PRESENT TEST DATA

Fruchter[2] says,

> As is true of all scientific reporting, it is desirable to give a complete account of a factorial study so that the results can be verified and the computations checked.

The same author[3] has listed some items that should be included in a complete account of a factorial study. In giving an account of the present study, all the suggestions that

---

1 Thurstone, L.L.,"Multiple-Factor Analysis", University of Chicago Press, Chicago, 1947, p. 152.

2 Fruchter, Benjamin, Op.Cit., p.153.

3 Fruchter, Benjamin, Op.Cit., pp. 153-154.

are appropriate to this study are carried out.

THE MEASURES

Though the application of the Wherry-Doolittle method of test selection suggests that the test III - Interest in Profession should be discarded and not be added in the final battery, the test is retained in the battery at least for the purpose of factor analysis. (The two studies of the tests are made separately and the results obtained from one study are not taken into consideration while the other study is made. This is a tentative arrangement. Final conclusion is arrived at from the combined results of both the studies).

The tests to be factor-analysed are:

(i) Attitude Towards Children.

(ii) Mental Ability.

(iii) Interest in Profession.

(iv) Adaptability.

(v) Professional Information.

The information regarding the scoring formulas, reliabilities, conditions of administration, distributions, means, standard deviations and measures of skewness and kurtosis is supplied at appropriate places in earlier pages.

THE SAMPLE

The same sample (N = 530) that was used for standardisation of the test was used for factor-analysis also.

THE CORRELATION MATRIX

The inter-correlations of five tests were found out. In each case the product-moment coefficient of correlation, r, was computed. The correlation matrix of the five variables was then prepared. It is shown in the following table.

TABLE NO. 76

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX
(N = 530)

Sr. No.	Test	I	II	III	IV	V
1	Attitude Towards Children	-	0.212	0.201	0.320	0.235
2	Mental Ability	0.212	-	0.129	0.222	0.214
3	Interest in Profession	0.201	0.129	-	0.186	0.151
4	Adaptability	0.320	0.222	0.186	-	0.160
5	Professional Information	0.235	0.214	0.151	0.160	-
		0.968	0.777	0.667	0.888	0.760

THE ANALYSIS

As stated earlier, Thurstone's centroid method was applied to the present test data for the extraction of factors.

The detailed step by step process of factoring suggested by Guilford[1] was followed by the investigator. The details of the steps followed are not given here. Simply the different matrices obtained are given in the following pages.

STAGE I: EXTRACTION OF THE FIRST CENTROID FACTOR

One simple method of estimating the communality of a test is to guess it to be equal to the highest correlation of that test with any other variable in the correlation table. The highest correlation in each column has been inserted in the principal diagonal cell of table No. 77, in parentheses.

TABLE NO. 77

EXTRACTION OF THE FIRST CENTROID FACTOR FROM THE CORRELATION MATRIX

Test	I	II	III	IV	V	Check sum
I	(0.320)	0.212	0.201	0.320	0.235	1.288 ✓
II	0.212	(0.222)	0.129	0.222	0.214	0.999 ✓
III	0.201	0.129	(0.201)	0.186	0.151	0.868 ✓
IV	0.320	0.222	0.186	(0.320)	0.160	1.208 ✓
V	0.235	0.214	0.151	0.160	(0.235)	0.995 ✓
E	1.288	0.999	0.868	1.208	0.995	5.358 $= \Sigma r = T$
$mE = a_1$	0.5564	0.4316	0.3750	0.5219	0.4299	2.3148 ✓ $\sqrt{T} = 2.315$

$$\frac{1}{\sqrt{T}} = 0.432 = m$$

$$mT = 2.315$$ ✓

1 Guilford, J.P., Op.Cit., pp. 485-500.

The first factor matrix was, then, prepared. It is given in the following table.

TABLE NO. 78

FIRST FACTOR MATRIX

First factor loading		0.5564	0.4316	0.3750	0.5219	0.4299
		I	II	III	IV	V
0.5564	I	0.3096	0.2401	0.2087	0.2904	0.2392
0.4316	II	0.2401	0.1863	0.1619	0.2253	0.1855
0.3750	III	0.2087	0.1619	0.1406	0.1957	0.1612
0.5219	IV	0.2904	0.2253	0.1957	0.2724	0.2244
0.4299	V	0.2392	0.1855	0.1612	0.2244	0.1848

STAGE II: COMPUTATION OF THE FIRST FACTOR RESIDUALS

The next step in the analysis was, then, to prepare the first residual correlation matrix. It is shown in the table on the next page.

TABLE NO. 79

FIRST RESIDUAL CORRELATION MATRIX

	I	II	III	IV	V	Check sum
I	(0.0104)	-0.0281	-0.0077	0.0296	-0.0042	0.0000
II	-0.0281	(0.0357)	-0.0329	-0.0033	0.0285	-0.0001
III	-0.0077	-0.0329	(0.0604)	-0.0097	-0.0102	-0.0001
IV	0.0296	-0.0033	-0.0097	(0.0476)	-0.0644	-0.0002
V	-0.0042	-0.0285	-0.0102	-0.0644	(0.0502)	-0.0001
Σ	0.0000	-0.0001	-0.0001	-0.0002	-0.0001	-0.0005

It is necessary at this stage to decide whether to proceed further for extracting the second centroid factor.

Fruchter[1] gives the following formula to find out the maximum number of factors which can be uniquely determined by n variables.

The formula is:

$$r = \frac{2n + 1 - \sqrt{8n + 1}}{2}$$

Where r = number of factors

n = number of variables.

In the present study, there are five variables. Therefore, the number of factors that can be expected is:

1 Fruchter, Benjamin, Op.Cit., pp. 68-69.

$$r = \frac{2 \times 5 + 1 - \sqrt{8 \times 5 + 1}}{2}$$

$$= \frac{10 + 1 - \sqrt{41}}{2}$$

$$= \frac{11 - 6.4}{2}$$

$$= \frac{4.6}{2}$$

$$= 2.3 \qquad \text{i.e. } 2 \text{ (Two)}$$

This suggests the possibility of the presence of a second centroid factor.

STAGE III: EXTRACTION OF THE SECOND CENTROID FACTOR

Variables I, III & IV were reflected so as to maximize the positive sum of the table of residual correlations (excluding the values in the principal diagonal).

TABLE NO. 80

EXTRACTION OF THE SECOND CENTROID FACTOR FROM THE FIRST RESIDUAL CORRELATION MATRIX IN WHICH VARIABLES I, III & IV ARE REFLECTED

Test	I	II	III	IV	V	Check sum
I	(0.0296)	0.0281	-0.0077	0.0296	0.0042	0.0838
II	0.0281	(0.0329)	0.0329	0.0033	0.0285	0.1257
III	-0.0077	0.0329	(0.0329)	-0.0097	0.0102	0.0586
IV	0.0296	0.0033	-0.0097	(0.0644)	0.0644	0.1520
V	0.0042	0.0285	0.0102	0.0644	(0.0644)	0.1717
E	0.0838	0.1257	0.0586	0.1520	0.1717	0.5918 = $\Sigma r = T$
$mE=a_2$	0.1089	0.1634	0.0762	0.1976	0.2232	0.7693 $\sqrt{T}=0.7693$

$$\frac{1}{\sqrt{T}} = 1.3 = m$$

$$mT = 0.7693$$

CRITERIA FOR SIGNIFICANT FACTORS

It should be tested whether the second centroid factor obtained is significant. There are no exact criteria for testing the significance of the factor obtained.

Vernon, as quoted by Fruchter[1], has listed as many as twenty five criteria for this purpose.

In the present study Humphrey's Rule[2] is applied.

1 Fruchter, Benjamin, Op.Cit., foot note on p. 77.

2 Ibid., pp. 79-80.

This criterion takes into account N, the size of the sample, and is dependent on the loadings of only two variables (which should be sufficient to establish a factor) rather than on the entire matrix.

The rule is:

(i) The product of the two highest factor loadings is found out. Here it is:

$$0.2232 \times 0.1976 = 0.044$$

(ii) The standard error of a correlation coefficient of zero, for the type of correlation and size of sample being used, is found out. It is

$\frac{1}{\sqrt{N}}$ for the Pearson product-moment r.

$$\therefore \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{530}} = \frac{1}{23.0217} = 0.043$$

$\therefore$ twice of it, is $= 0.086$.

(iii) If the product (0.044) found in step (i) does not exceed twice the standard error (0.086) found in step (ii), the factor is probably not significant.

Here, the product of two highest factor loadings is less than twice the standard error, and hence the factor concerned is not present.

Therefore, there is only one factor present, in the five variables studied.

CORRELATION MATRICES FOR DIFFERENT UNIT-SAMPLES

The product-moment coefficients of correlation between tests for all the different unit-samples were also calculated. The factor-analysis was applied to the data obtained from unit-sample F. Here also only one factor was extracted. The correlation matrix for each unit-sample was prepared. These matrices are given below. Comparing all the six matrices, it can be said that they are not similar to a very great extent but if all were factor analysed, one factor only, in each case would have been obtained.

TABLE NO. 81

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE - A - N = 100)

Test	I	II	III	IV	V
I	-	0.160	0.076	0.180	0.146
II	0.160	-	0.045	0.200	0.021
III	0.076	0.045	-	0.294	0.150
IV	0.180	0.200	0.294	-	0.076
V	0.146	0.021	0.150	0.076	-
	0.562	0.426	0.565	0.750	0.393

TABLE NO. 82

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE-B) (N = 74)

Test	I	II	III	IV	V
I	-	0.055	0.188	0.456	0.199
II	0.055	-	0.266	0.169	0.283
III	0.188	0.266	-	0.011	0.023
IV	0.456	0.169	0.011	-	0.349
V	0.199	0.283	0.023	0.349	-
	0.898	0.773	0.488	0.985	0.854

TABLE NO. 83

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE-C) (N = 66)

Test	I	II	III	IV	V
I	-	0.108	0.009	0.494	0.179
II	0.108	-	0.358	0.270	0.194
III	0.009	0.358	-	0.087	0.119
IV	0.494	0.270	0.087	-	0.000
V	0.179	0.194	0.119	0.000	-
	0.790	0.930	0.573	0.851	0.492

TABLE NO. 84

INTERCORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE-D) (N = 94)

Test	I	II	III	IV	V
I	-	0.206	0.410	0.197	0.386
II	0.206	-	0.119	0.136	0.156
III	0.410	0.119	-	0.220	0.282
IV	0.197	0.136	0.220	-	0.182
V	0.386	0.156	0.282	0.182	-
	1.199	0.617	1.031	0.735	1.006

TABLE NO. 85

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE-E) (N = 78)

Test	I	II	III	IV	V
I	-	0.105	0.075	0.255	0.368
II	0.105	-	0.190	0.164	0.484
III	0.075	0.190	-	0.130	0.285
IV	0.255	0.164	0.130	-	0.249
V	0.368	0.484	0.285	0.249	-
	0.803	0.943	0.680	0.798	1.386

TABLE NO. 86

INTER-CORRELATIONS OF FIVE TESTS, ILLUSTRATING THE CONDITION OF A SIMPLE PROPORTIONALITY IN A CORRELATION MATRIX (UNIT-SAMPLE-F) (N = 118)

Test	I	II	III	IV	V
I	-	0.367	0.164	0.273	0.313
II	0.367	-	0.116	0.205	0.183
III	0.164	0.116	-	0.185	0.044
IV	0.273	0.205	0.185	-	0.229
V	0.313	0.183	0.044	0.229	-
	1.117	0.871	0.509	0.892	0.769

The factor analysis is applied to reduce the number of variables in a test-battery. According to Guilford[1], for this purpose, almost any method of factor analysis will do, with or without rotations of axes.

CONCLUSION

The factor analysis shows that there is one factor present in the five variables studied. The factor thus extracted is a common factor which has appreciable loadings on all the tests in a battery. It may, therefore, be called a general factor for the battery.

There are two possibilities regarding the interpretation of this general factor. Either this factor is identical

1 Guilford, J.P., Op.Cit., p. 522.

with 'G' factor found in tests of general intelligence or that it is a general factor called "APTITUDE FOR TEACHING" and different from 'G'.

A remark from Froehlich and Hoyt[1] is significant in concluding this chapter. They say,

> It is important to note that findings from factor analysis of tests are based almost solely on statistical procedures. So these findings should be used only after they have been checked with common-sense analysis.

---

1 Froehlick, C.P. & Hoyt, K.B., "Guidance Testing", Science Research Associates, Inc., Chicago, 1959, p. 65.