

ment test for foreign language aptitude. When an achievement test was used, the correlation jumped to .64.<sup>21</sup>

From our examination of the problem of prediction, we draw the conclusion that a fruitful point of attack is through the substitution of a more reliable and therefore more predictable measure of achievement. This paper has presented data which definitely demonstrates that a pencil-and-paper evaluation instrument such as the Sophomore Culture test is more predictable than the time-honored grade criterion. But it would be foolhardy indeed for the authors to take the next step, that of advocating that this attribute alone justified its substitution for honor-point ratio. This decision lies within the province of the educational administrator. He must decide whether more accurate prediction — a *sine qua non* of all efficient admissions policies — plus the Culture test's degree of relevance is sufficient to outweigh those desirable qualities which may still be claimed for the traditional marking system. In short, he must decide whether this new criterion is more acceptable than the old.

A final word for research. The Sophomore Culture test, in common with other achievement tests, largely measures recall of information.<sup>22</sup> That information is only one phase of education must be recognized. Other components of cultural growth — attitudes, values, motivations, goals, and affective experience — must be measured by other instruments. It is hoped that in the not-too-distant future these important outcomes of education can be appraised with sufficient accuracy so that we may know how well the American college functions as the vehicle of culture.

<sup>21</sup>J. B. Tharp, "Sectioning Classes in Romance Languages," *Modern Language Journal*, XII (1927), 95-114.

<sup>22</sup>B. E. Cureton, "Evaluation or Guidance—A Report of the 1939 Sophomore Testing Program," *Journal of Experimental Education*, VIII (1940), 308-40.

6462  
BLAKEI

## A FACTOR ANALYSIS OF A NON-VERBAL REASONING TEST

ROBERT I. BLAKEY  
Social Security Board

SOME time ago Dr. Andrew W. Brown and the author constructed a "Non-Verbal Reasoning Test" for use at the high school level. A preliminary report of its construction is being published by *The Journal of Educational Psychology*. The present article concerns itself with the results of a factor analysis of the intercorrelations between the subtests rather than with the actual standardization of the test.

The test was constructed with the idea that it should measure in a non-verbal manner the higher intellectual processes of comprehension, mental alertness, deductive reasoning, inductive reasoning, and spatial relations or analysis. The primary purpose of this study is to isolate and identify any common factors present and to compare them with the expected factors.

Other problems which may be considered in the light of the factor analysis are: (a) a comparison of the factorial composition of tests which are variations of Thurstone's tests with the factorial composition, as determined by Thurstone, of the tests he used; (b) a reconsideration of the perennial problem of the existence of a general factor of mental ability; (c) the comparison of the factors found in this group of tests with factors found in analyses of other tests; (d) a further examination of various methods of ascertaining the number of factors which should be taken out of a correlation matrix.

All tests are time-limit tests and were introduced by fore-exercises which were explained by the examiner. They were presented in the order listed.

1. *Manikin*—a page of pied figures of little men. The figures are simple line drawings with variations in the positions of arms and legs. The problem is to draw a ring around each manikin which is exactly like a model at the top of the page. It was thought that this test might be saturated with the Perceptual Speed factor. The Spearman-Brown corrected reliability is .81.

2. *Identical Patterns*—12 rows of patterns formed by overlapping geometrical forms. The first pattern of each row is separated from the others by a heavy vertical line. The patterns are in 12 variations each composed of two circles and two right triangles. The same size forms are used in each variation, the differences being due to relative positions of the components and whether the forms are solid or dotted lines. Each row contains one or more patterns exactly like the first one in the row, and the problem is to place a mark under each pattern which is exactly like the first one in its respective row. It was thought that this test would be a variation of Thurstone's *Identical Forms* test and consequently loaded with the Perceptual Speed factor. The Spearman-Brown corrected reliability is .98.

3. *Fitting Parts*—each item consists of a solid black geometrical form, which has been cut into three parts, and four outlined figures, one of which is the same size and shape as the black figure which was cut. The problem is to indicate that one of the outline forms into which the solid black pieces could be made to fit exactly. Discrimination of both size and shape is involved for each item. It was thought that possibly the factor Visualization or Space was involved in the solution of this test. The Spearman-Brown corrected reliability of the 12-item test is .47.

4. *Opposite Sides*—each item consists of three drawings identical in size and shape. The problem is to select the drawing in each item which is a mirror image of the other

two drawings. Each drawing is a little pennant the shape of a non-isosceles right triangle and may be rotated in any position. It was thought that possibly Space and Induction might be used in the solution of this test. There is no really parallel form to this test although the idea was adopted from Thurstone's *Flags* test. The Spearman-Brown corrected reliability is .88.

5. *Code*—a code consisting of eight boxes divided in half is placed at the top of the test. Each box has a unique group of squares and circles in the top half and an unusual group of triangles in the bottom half. Below the "code" are five rows of the little boxes, some exactly like the boxes in the code and some with incorrect pairing of the symbols. The problem is to place a line under each box which is different from the code. It was thought that the test might contain the Perceptual factor. The Spearman-Brown corrected reliability is .96.

6. *Circle Grouping*—each item consists of four boxes containing little groups of circles. The grouping varies from box to box. One circle in each of the first three boxes is blackened according to a system. The problem is to discover that system and apply it in blackening a circle in the fourth box. It was thought that possibly Induction would be involved in solving this test. The Spearman-Brown corrected reliability for the 12-item test is .98.

7. *Form Series*—this test is the usual series type with only three meaningless forms used in combination. One figure in each row is omitted and a blank inserted. The problem is to indicate which form belongs in the blank. It was thought that Deductive Reasoning or Inductive Reasoning would be involved in the solution of this test. The corrected Spearman-Brown reliability of the 22-item test is .86.

8. *Circle Reasoning*—a variation of the *Marks* test used by Thurstone as a measure of Inductive Reasoning. There are five rows of groups of circles and dashes. The grouping changes from row to row. One circle in each of the first

four rows is blackened according to a rule. The problem is to find the rule and apply it in blackening a circle in the fifth row. It was assumed that this test would contain Induction. The corrected reliability is .94.

9. *Form Relations*—this test is a parallel form of Thurstone's *Pattern Analogies* test. The problem is to indicate one of five choices which bears the same relation to the third figure as the second bears to the first. Inductive Reasoning or Deductive Reasoning was assumed to be necessary for the solution of this test. The corrected reliability is .97.

10. *Form Reasoning*—at the top of the test is a table showing how any two of seven forms could be combined to equal another one of the seven. Each item consists of three of the forms in a row. The task is to combine the first two forms according to the table and then combine the resulting form with the third to equal another form, the final result to be indicated by underlining one of five choices. It was thought that possibly Deductive Reasoning would be used to solve these problems. The Spearman-Brown corrected reliability for the 12-item test is .98.

*The Subjects*

The subjects were 286 high school pupils from a school in a suburb of Chicago. All tests were given by two experienced examiners in a well-lighted room. All tests were administered in one 40-minute period. Eighty per cent of the whole group was between 15 and 18 years of age. The mean Otis I.Q. was 114. About 54 per cent of the group were boys. No sex difference was found for combined scores on the whole test. No grade difference was statistically significant. The correlation of total test score with chronological age was  $-.13$  for the age range of this group.

*The Factor Analysis*

The table of intercorrelations (Table 1) was computed with the aid of *Computing Diagrams for the Tetrachoric Correlation Coefficient* (2). Correlations obtained in this

manner are considered by Thurstone (6, p.58) to be applicable to factor analysis. In effect the scores are normalized in the process of correlation.

The factors (Table 2) were extracted by the Thurstone centroid methods. Here the problem of the number of factors

*(X) Lower number*

TABLE 1  
INTERCORRELATIONS OF TESTS

Variable	1	2	3	4	5	6	7	8	9	10
Manikin .....		.24	.27	.24	.38	.19	.13	.19	.22	.19
Identical Patterns....	.24		.08	.17	.22	.46	.16	.15	.33	.24
Fitting Parts.....	.27	.08		.17	.22	.20	.13	.10	.20	.22
Opposite Sides.....	.24	.17	.17		.15	.25	.38	.32	.39	.31
Code .....	.38	.22	.22	.15		.26	.22	.25	.35	.38
Circle Grouping.....	.19	.46	.20	.25	.26		.48	.50	.53	.49
Form Series.....	.13	.16	.13	.38	.22	.48		.35	.52	.54
Circle Reasoning.....	.19	.15	.10	.32	.25	.50	.35		.55	.38
Form Relations.....	.22	.33	.20	.39	.35	.53	.52	.55		.40
Form Reasoning.....	.19	.24	.22	.31	.38	.49	.54	.38	.40	

TABLE 2  
CENTROID MATRIX (F)

Variable	Code No.	I	II	Factors III	IV	V
Manikin .....	1	.438	-.435	-.183	-.083	-.069
Identical Patterns.....	2	.452	-.141	.274	.263	-.200
Fitting Parts.....	3	.335	-.212	-.101	-.140	.087
Opposite Sides.....	4	.499	.100	-.163	-.112	-.242
Code .....	5	.506	-.297	-.055	-.079	.130
Circle Grouping.....	6	.701	.138	.296	.272	.109
Form Series.....	7	.622	.377	.110	-.274	-.117
Circle Reasoning.....	8	.602	.281	-.252	.238	.205
Form Relations.....	9	.728	.181	-.154	.177	-.093
Form Reasoning.....	10	.665	.119	.166	-.251	.239

appeared. Two methods of determining the number of factors had been tried by the author (1) previously with some degree of success. One of these, Tucker's empirical criterion, gave negative results in the present case. The other, Coombs' criterion (3) postulates that in a 10-variable problem, the last factor of value will leave a table of residuals which, when signs are changed, will contain more than 31 negative entries with a standard error of five. Table 3 shows the application of Coombs' criterion to this analysis:

TABLE 3  
COOMBS' CRITERION

Factor	Negatives
1	24
2	33
3	24
4	28
5	35

It was obvious from the number of relatively large residuals remaining in the table after the second factor was extracted that there were more than two factors in the table. This was borne out in the subsequent analysis, which was carried to five factors. The indication that the fifth factor was the last one of value seems to have been verified in the analysis. The standard deviation of the fifth factor residuals before sign change is .028, which is considerably smaller than the standard error of a zero correlation for a population of 286.

For the rotation of factors in order to secure bounding hyperplanes, Thurstone's method of lengthened vectors was used (4). The criteria of maximizing the number of zeros and rotating to a postulated positive manifold were the determiners for direction of rotation. Seven rotations were necessary and a "clean-up" rotation with actual length vectors was made. The rotated factorial matrix is given in Table 4. The rotational matrix of direction cosines is given in Table 5. The intercorrelations between the rotated factors are presented in Table 6.

TABLE 4  
ROTATED FACTORIAL MATRIX (FA)

Variable	Code No.	Factor				
		A	B	C	D	E
Manikin	1	.582	.075	.004	.192	.054
Identical Patterns	2	.092	.547	-.016	.239	.014
Fitting Parts	3	.345	-.041	-.009	.265	.005
Opposite Sides	4	.152	.028	.160	.313	.408
Code	5	.440	.067	.004	.394	-.062
Circle Grouping	6	-.076	.436	.161	.641	-.073
Form Series	7	-.141	.040	.016	.639	.453
Circle Reasoning	8	-.010	-.046	.561	.518	.026
Form Relations	9	.076	.162	.415	.507	.244
Form Reasoning	10	.080	.021	-.053	.766	.071

TABLE 5  
TRANSFORMATION MATRIX (A)

Centroid Axis	Reference Vector				
	A	B	C	D	E
I	.287	.247	.207	.801	.203
II	-.859	-.243	.361	.248	.379
III	-.380	.671	-.690	.234	-.224
IV	-.184	.521	.582	-.257	-.436
V	.033	-.399	.111	.421	-.759

TABLE 6  
CORRELATIONS BETWEEN NORMALS TO THE PLANES (A' A)

Plane	Plane				
	A	B	C	D	E
A	1.000				
B	-.084	1.001			
C	-.092	-.241	1.000		
D	-.011	-.007	-.009	1.001	
E	-.127	-.117	-.005	-.003	1.001

Even a cursory glance at the rotated matrix will show that the factorial composition of the tests is not so simple as had been hoped for.

Factor "A" has three variables with significant projections and all the others are essentially zero. These are:

1. *Manikin* ..... .58
3. *Fitting Parts* ..... .35
5. *Code* ..... .44

Either one of two interpretations could be placed on this factor. It might be considered to be Space as has been described by Thurstone (6), the author (1), and others. Under this interpretation it would seem that the grasping of spatial relations of the arms and legs of the manikins was of more importance than the quick perception of small differences. It would appear also, that the quick comparison of the code with the stimuli in the *Code* test was not so important in solving the problem as the grasping of the relationship between the two halves of the individual elements.

The other interpretation which could be placed on this factor is that it is Perceptual Speed, or rather mental alert-

ness, as distinguished from Perceptual Discrimination. Under this interpretation the ability would involve the quick change of response from item to item with only the simplest discrimination necessary. Thurstone's factor "9" in his study of Hyde Park High School in Chicago seems to have some of the characteristics of factor "A" (5). In this case, the test *Scattered "x"s* had the highest loading. The *Manikin* test has the simplest discrimination level and the *Fitting Parts* test the most complex of those listed. The author prefers this latter interpretation.

Factor "B" has two tests which have significant loadings:

2. <i>Identical Patterns</i> .....	.55
6. <i>Circle Grouping</i> .....	.44

It seems obvious that this factor corresponds to Thurstone's (6) Perceptual Speed factor, but we shall call it Perceptual Discrimination to distinguish it from factor "A." The difference here is that the emphasis is on analytic perception in which a fine discrimination must be made rather than on speedy response to a simple stimulus. Speed is of importance, but in the subjects used the differences in the mental process of perceptual discrimination will contribute more to performance variance than will simple speed.

At first glance it seems surprising that *Circle Grouping* is high on this factor. However, a careful subjective analysis of the test will indicate that the problems involved are more those of perceptual discrimination than of induction. The figures are complex but the rules to be brought out are simple. For example, one of the items has the middle dot blackened in a group of three, which is apparent even at a glance, so that the problem resolves into finding the correct group in the response square. This takes a discriminatory ability evidently slightly below that required for *Identical Patterns*.

Factor "C" has two variables with significant projections:

8. <i>Circle Reasoning</i> .....	.56
9. <i>Form Relations</i> .....	.42

Both of these tests are variations of tests used by Thurstone in his studies of the primary mental abilities and have been interpreted to contain Induction, or Inductive Reasoning. This interpretation is suitable in the present case. The apparent paradox that test 8 contains Induction while test 6, in which a supposedly similar function is involved, does not may be resolved when an inspection is made of the tests themselves. The primary problem in test 8 is to find a rule by which the problem may be solved while in test 6 the main problem, as has been said before, is to find the response group rather than the rule.

Factor "D" is an orthogonal factor which was set up by making its normal perpendicular to the normals of all the other planes. This was necessary as one dimension of the five-dimensional system could not be identified by a bounding hyperplane because of lack of variables with zero projections in that dimension. It is the same type of problem as was encountered by the author in a former study (1).

All the variables have projections on this factor which are probably significant. The relative amount of projection seems to increase with the greater complexity of the mental function involved. The tenth test, *Form Reasoning*, which involves the synthesis of geometrical figures according to established rules (not unlike arithmetic), has much the highest saturation of the factor.

The obvious comment, and one that must be reckoned with, is that this factor represents "general intelligence," or Spearman's factor "g." As has been said before, there is nothing in the Thurstone method of analysis which denies that such a general factor exists or implies that it would not show up if present. However, in regard to the nature of the present factor, there can be little doubt that it is "general" for this battery of tests and is not an effect of maturation or lack of differentiation of ability due to the youth of the subjects. What it is called—*comprehension, understanding, mental efficiency, or intelligence*—is beside the point. Due to the popular misconceptions and scientific vagueness of the last

term, it probably would be better to adopt some other name.

It should be understood that the author is of the opinion that the above-mentioned effect of an augmented general factor due to lack of maturation is applicable to situations in which the subjects are immature, but that such a factor does not account for appreciable distortion in the present case. It is not denied that such a general factor is present in tests given to children, but it seems probable that the general factor, if it exists in such a case, is unduly emphasized by the maturation curves of the abilities.

Another interpretation which might be placed on factor "D" is that it is Deductive Reasoning, which in each test requires that the subject must base his conclusions or responses on certain facts which are presented in the test item. However, this is probably another aspect of the foregoing discussion.

Factor "E" has significant loadings for two tests and a possibly significant loading for a third:

4. <i>Opposite Sides</i> .....	.41
7. <i>Form Series</i> .....	.45
9. <i>Form Relations</i> .....	.24

This factor apparently corresponds with none of the factors previously identified by Thurstone and his associates. However, it may possibly represent Deductive Reasoning as "series" tests have been found by Thurstone (5) to contain a component of Deductive Reasoning. The same is true of the form relations type of test. The relationship of the *Opposite Sides* test to such an interpretation is not immediately apparent. Assuming that one might consider two figures in each item of the *Opposite Sides* test as facts to be compared and from which a conclusion might be drawn concerning the third figure, i.e., whether it is different from the first two or like one of them, then it might be thought to involve Deduction. In the *Form Series* test the symbols presented are facts from which a conclusion must be drawn concerning the missing figure. The conclusion is definitely limited to three alternatives

each of which might be tried in turn. In the *Form Relations* test the problem might be approached by trying to find the rule involved, which would be Induction, or by substituting the possible answers one at a time and testing the resulting equation. This latter process might be considered to be Deductive Reasoning and insofar as it were used would cause the test to show a loading on the Deduction factor.

No definite conclusion can be made as to the identity of Factor "E," but tentatively it may be called Deductive Reasoning.

Despite the fact that the factorial composition of some of the tests varies somewhat from what was originally supposed, it seems that the tests, as a group, do measure some of the higher mental processes of reasoning. From amount of projection on the general factor, it would seem that the tests saturated with Perceptual Speed are the poorest measures of the higher intellectual processes. It would appear that test number 9, *Form Relations*, which has significant projections on three factors, is probably the best general test of all the reasoning processes. Test 10, *Form Reasoning*, is the best test of the general factor which might be considered to be synonymous with comprehension or mental efficiency or intelligence. The test, *Identical Patterns*, seems to be saturated with the factor Perceptual Discrimination, which is interpreted quite similarly to Thurstone's factor of Perceptual Speed, and is consistent with Thurstone's (5) test of *Identical Forms*, which is parallel in process. The test, *Circle Reasoning*, a variation of Thurstone's (5) *Marks* test, is similar in factorial composition to the latter. The *Form Relations* test seems to have a heterogeneous factorial makeup, as was also found by Thurstone (6).

The factors identified seem to be consistent with those identified by Thurstone (6,5) except for the general factor. It is necessary to investigate these tests in a larger battery before an interpretation can be adequately applied to the general factor. This factor has some characteristics similar to those found by the author (1) in factor "D" in a "Reanalysis

of a Test of the Theory of Two Factors." The factor Perceptual Speed also seems similar to the factor "C" in the latter study.

The factors have been found to be practically uncorrelated, the highest correlation, that between factors "B" and "C," being only 14 degrees off orthogonality. This is probably within chance variation and no significance is attached to it.

## REFERENCES

1. Blakey, R. I. "A Reanalysis of a Test of the Theory of Two Factors," *Psychometrika*, II (1940), 121-36.
2. Chesire, L., Saffir, M., and Thurstone, L. L. *Computing Diagrams for the Tetrachoric Correlation Coefficient*. Chicago: University of Chicago Press, 1933. 59 pages.
3. Coombs, Clyde. Unpublished paper read before the American Psychological Association, September, 1940.
4. Thurstone, L. L. "A New Rotational Method in Factor Analysis," *Psychometrika*, III (1938), 199-218.
5. Thurstone, L. L. "Experimental Study of Simple Structure," *Psychometrika*, II (1940), 153-68.
6. Thurstone, L. L. *Primary Mental Abilities*. Chicago: University of Chicago Press, 1938. 121 pages.