PSYCHOMETRIKA—VOL. 26, NO. 4
DECEMBER, 1961

AN EMPIRICAL STUDY OF THE FACTOR ANALYSIS STABILITY HYPOTHESIS*

HAROLD P. BECHTOLDT
UNIVERSITY OF IOWA

Note is taken of four related sources of confusion as to the usefulness of Thurstone's factor analysis model and of their resolutions. One resolution uses Tucker's distinction between exploratory and confirmatory analyses. Eight analyses of two sets of data demonstrate the procedures and results of a confirmatory study with statistical tests of some, but not all, relevant hypotheses in an investigation of the stability (invariance) hypothesis. The empirical results provide estimates, as substitutes for unavailable sampling formulations, of effects of variation in diagonal values, in method of factoring, and in samples of cases. Implications of these results are discussed.

It has been said that a work of art should provoke favorable or unfavorable reactions, and that a scientific theory should lead to further empirical and theoretical work. Factor analysis, which has been called both an art and a scientific approach to the study of individual differences, certainly has evoked strong emotional responses as well as extensive empirical and theoretical studies. However, neither reaction has succeeded in clarifying either the role of factor analysis or the appraisal of its usefulness as a research technique. One might say confusion, if not chaos, is the norm in this field.

In the development of a science of psychology, confusion about the usefulness of a set of procedures such as those of factor analysis should be a matter of great concern. This paper takes the position that the present confusion stems, in part, from disagreements as to definitions of terms or concepts, and, in part, from failures to make certain analytical distinctions. A recent symposium on the "Future of factor analysis" [42] exemplifies several of these semantic and analytical confusions. The objectives of the present paper are threefold: (i) to call attention to these sources of confusion, to some of their implications, and to procedures for resolving the confusion; (ii) to demonstrate a type of factor analysis in which the usefulness of some hypotheses related to the stability or invariance of factor analysis data can be evaluated; and (iii) to provide, for several factoring procedures, empirical estimates of the sampling variation in objectively determined oblique simple

*The computational costs of this study were defrayed, in part, by a research small grant M-1922 from the National Institute of Health, and, in part, by support under project 176-0002 by the University of Iowa Computing Center, Dr. J. P. Dolch, Director. The assistance of Dr. Kern Dickman and Mr. Leonard Wevrick of the University of Illinois and of Mr. Norman Luther of the University of Iowa in handling the computing problems is gratefully acknowledged.

structure values based on data from two samples. Since available solutions to some statistical problems associated with factor analyses are impractical, such empirical estimates are useful but limited substitutes for the desired analytical formulations.

Sources of Confusion

One of the most insidious and ubiquitous sources of confusion is the ambiguity of the term *factor analysis*. That different models are classed under this single term is well known, but discussions of the objectives, techniques, and results obtained often do not make clear the specific model under consideration. For example, the principal components model must be distinguished from factor analysis models. The several factor analysis models in turn are differentiated on the basis of the postulation of a general factor, the acceptance of the necessity for rotation, and the criteria for the final solution (orthogonal vs oblique axes, simple structure, etc.). Statements appropriate to one model often do not apply to another even though certain transformations from one model to another may exist.

The present discussion deals with the model as formulated by Thurstone [52, 54, 55] and extended or modified by the work of Anderson and Rubin [2], Bargmann [3, 4], Guttman [29, 30], Howe [36], Koopmans and Reiersöl [38], Lawley [39], Rao [46], Reiersöl [48], and Tucker [62, 63]. These papers indicate the relevance to factor analysis theory of the problems concerned with (i) the existence of the model (solvability), (ii) the identification of the parameters (uniqueness), (iii) the determination of the number of factors (rank), (iv) the criteria for rotational transformations, and (v) the test of the hypothesis that the model fits a set of data. The acceptability of Thurstone's formulation involving oblique simple structure with common and unique factors requires consideration of solutions to these problems, several of which Thurstone explicitly recognized ([54]; [55], pp. v–xiv). Unfortunately, the extensions and clarifications of Thurstone's earlier work in the articles noted above have not been considered in recent articles critical of this factor analysis model [32, 33, 42, 58, 59, 60, 66, 67]. A minor but frustrating related source of confusion is the introduction of different names for the operations and concepts of Thurstone's model [60]. Furthermore, the implications of these papers have not been adequately considered by some proponents [17, 26] who use this model of factor analysis with test data. The result has been a proliferation of irrelevant and unacceptable arguments as to the usefulness of Thurstone's method as well as that of any other model of factor analysis.

A second source of confusion derives from failures to make or to maintain distinctions related to the objectives of an investigation. Such differences in purpose exist between the type of factor analysis which Tucker [63] calls *exploratory* as opposed to the type that Tucker calls *confirmatory*.

Basically his distinction depends upon the amount of information and of precision of knowledge in an area. The exploratory factor analysis, being the first, is used to generate hypotheses while a confirmatory factor analysis is designed subsequently to test these hypotheses. It is generally accepted as a principle of hypothesis testing that the same set of data cannot be used both for inventing or generating hypotheses and for evaluating the usefulness of the hypotheses. It is, of course, conceivable that an initial analysis could be sufficiently precise to permit the use of the term *confirmatory* and to permit the testing of certain hypotheses.

The purpose of an exploratory analysis, as stated by Thurstone, is "... to discover the principal dimensions or categories ... and to indicate the directions along which they may be studied by experimental laboratory methods" ([54], p. 189; [55]). The principal dimensions are discovered by the appearance of trans-situational response consistencies defined by the operations of factor analysis discussed in detail by Thurstone [55] and Tucker [61, 63], for example. This objective also can be expressed as the development of definitions of new composite variables and as the invention of hypotheses involving such variables [8]. In either type of activity the creative, artistic judgment of an investigator is as relevant in an exploratory factor analysis as in other creative endeavors; but the prior compulsions of the investigator for orthogonality or for a general factor, for example, will also be represented in his judgments and formulations [18]. The reader of the factor analysis literature should recognize that different compulsions lead to different results, i.e., to factors differently defined. Factors from two or more studies logically are not the same factors unless the defining operations including the reference tests and factoring procedures are the same. They may be similar factors or even parallel factors, provided definitions of such terms are specified, as Gulliksen does for parallel tests [27].

Since the initial formulation or invention of a variable or of a hypothesis cannot be useful by fiat, subsequent research to evaluate this usefulness is necessary. For this latter purpose, one may compare the empirical results using new samples from the same population with those obtained in the initial investigation; this procedure checks the stability or the invariance of the factor pattern [55]. Another stability question deals with the consistency of the empirical relations among modified or improved reference variables with those observed with the initial or unimproved variables, i.e., invariance under changes in the stimulus-response features of the task [6, 26]. Other hypotheses formulated in the initial exploratory study may deal with the number of significant factors and the location of specified zero factor loadings. For the investigation of such questions, subsequent confirmatory factor analyses on new samples of cases would be appropriate. Completely objective techniques for the conduct of such studies, if they are properly designed, are available together with some of the desired statistical tests. The design of such confirmatory studies is not a matter of guesses or hunches, nor can just any available table of correlations be used because specific and often testable hypotheses are involved. The problems of designing such studies have been considered repeatedly by Thurstone [52, 53, 54, 55] and by Tucker [63]; necessary and sufficient conditions for existence and uniqueness of solutions are indicated by Anderson and Rubin [2].

A third source of confusion in the literature is associated with the types of hypotheses that can be evaluated by factor analysis procedures. Some aspects of this distinction have been noted by Eysenck [21] and by Peel [44]. Only a few specific hypotheses of the many considered in the factor literature can be appropriately investigated with the conventional factor analysis models. For many hypotheses, a distinction is made, or implied, in the statements of the hypotheses between a set of reference variables and/or a set of treatment conditions as the independent variables on the one hand and the experimental or dependent variables being studied on the other. This distinction is associated with differences in status for these two classes of variables. For factor analyses, Thurstone specifically rejects this distinction between *independent and dependent* variables ([54]; [55], p. 59); *in fact*, the accepted principle in the several factor analysis models is that all variables are to be treated as coordinate or equal. Thus, for hypotheses expressing one or more variables as functions of one or more other variables, the usual factor analysis model is inappropriate (unless the modifications noted below are made). Such hypotheses include those dealing with the effects upon factor scores of variations in age, kind of instruction, amount of practice, drugs, or genetic history, for example. Other nonfactorial hypotheses include those dealing with factors as sources of variance in scores on tasks not used in the definitions of the factors. Both the distinction between independent and dependent variables and the introduction of approximations to part-whole correlations are points of issue in attempts to use a factor method for such hypotheses. In addition, the problem of communalities and the process of standardizing scores in computing correlations create further difficulties for between-group comparisons [47].

A fourth source of confusion arises from the description of factors as underlying causal variables which are not observable, which can only be inferred from the response consistencies, and which cannot be explicitly defined. This linguistic formulation involves a debated point in the philosophy of science, a point Bergmann [11, 12] calls the confusion between meaning and significance. A related argument is treated by Henrysson [34] in a discussion of explanatory factor analysis. The problem for factor analysis is that such unobservable variables cannot be directly studied as can other defined concepts. The factors cannot be investigated in the laboratory, for example, as suggested by Thurstone nor can the relations between factors and other variables be evaluated by nonfactorial methods, a procedure considered important by Thurstone [54, 55] as well as by most experimental psychologists. The factors are treated as existential hypotheses or almost as reified entities [20]. Brodbeck has made several pertinent points regarding this manner of speaking [13, 14]. And such writers as Anderson and Rubin [2], Koopmans and Reiersöl [38], and Rao [46] also have noted some of the difficulties associated with the unobservable characteristic of factors.

The fourth source of confusion can readily be resolved by using the results of an exploratory factor analysis and possibly of one or more confirmatory analyses to provide explicit objective definitions of the factors. These definitions will specify a factor as a definite function of observations on one or more designated reference variables. Such definitions are consistent with the existence of such factored tests as Thurstone's PMA battery [57] or of such sets of factor reference tests as the ETS Kit [24]. This resolution is consistent with Thurstone's statement of the objective of factor analysis; it also has several important linguistic and empirical implications. For example, the identification of factors as the same factor is not a problem [10] nor are procedures for defining a factor space common to two test batteries [64]. The third source of confusion can then be resolved by using explicitly defined factors as predictors in combination with the separation of independent and dependent variables in the analysis. With these two modifications of the factor analysis model, the several factor techniques can be shown to be ways to compute beta weights in the linear regression model. A convenient computing procedure uses the operations of the multiple-group method to project the dependent variables onto the space of the independent or factor variables. This relation follows directly from the early work of Holzinger and Harman [35] and Young and Householder [68]. In addition, explicitly defined factors can be used in nonfactorial experiments either as independent variables or as dependent variables. When a factor is explicitly defined without restricting the sample means and variances, the scores on the defined factors can be used as any distribution of test scores is used. The usefulness both of hypotheses involving factors and of proposed definitions of a factor then can be evaluated by the procedures regularly used for other hypotheses and other concepts.

Confirmatory Factor Analysis

The present study provides a demonstration of a confirmatory factor analysis conducted with a set of objective procedures in an investigation of the invariance of a simple structure solution, i.e., of the stability hypothesis. A series of questions related to the testable hypotheses of a confirmatory factor analysis are investigated; the relevance of these questions has been emphasized by Maxwell [41]. Answers to the questions were obtained from data on two samples of cases for a set of seventeen reference variables hypothesized on the basis of previous factor studies to be associated with a given number (six) of factors and with a specified set of zero and nonzero factor loadings. The following five questions are considered.

1. Is the hypothesis of two independent random samples from a single multivariate normal population tenable with reference to two sets of means and two variance-covariance matrices?

2. If the first hypothesis is tenable, can the set of 17 variables for each sample be considered as demonstrating some significant amount of dependency as defined by Bargmann ([4], pp. 43–68) (i.e., the rejection of the hypothesis of independence)?

3. If the first two hypotheses are tenable, does the degree of dependence (number of factors), as defined by the maximum likelihood or canonical correlation procedures for each sample, correspond to a hypothesized value—namely six?

4. If the first three hypotheses are tenable, does the factor pattern of zero and nonzero loadings for each sample as defined by an oblimax analytical rotation correspond to the hypothesized factor pattern for the results obtained from three factoring methods applied to the data, including one or more of four sets of estimated communalities?

5. Are the results of the simpler graphical (judgmental) rotational methods and of the multiple-group methods without rotation consonant with those obtained from other methods of analysis?

Data

The data are a portion of those originally collected by Thurstone and Thurstone ([56], ch. 3) for an analysis involving the hypothesis of seven primary mental abilities. Statistical tests of the relevant hypotheses were not then available. On the basis of the earlier analysis, the definition of the *Perceptual Speed* factor was judged by Thurstone to be inadequate, and it was therefore dropped from the present study of sampling effects. In addition, one of the variables for the *Memory* factor, the Figure Recognition test, was eliminated as being an unacceptable defining variable for the factor M. The 17 remaining tests were then considered as defining six primary mental abilities (PMA's) by six isolated constellations such that variations in the rotated factor loadings would provide a useful estimate of the sampling fluctuations for the statistics under investigation. Two samples of cases ($N = 212$ and $N = 213$) were formed by assigning each of 425 cases alternately to one or the other of two groups after the cases were thoroughly randomized. The original data in Tables 1, 2, and 9 were computed by Dorothy Case Bechtoldt in an unpublished study under the direction of L. L. Thurstone and L. R Tucker.

The list of 17 variables along with the means and standard deviations for these two samples as well as the hypothesized nonzero factor loadings for each variable are presented in Table 1. The location of each nonzero

TABLE 1

Seventeen Variables With Sample Means and Standard Deviations

Code No.	Name of Variable	Sample I (N=212) Mean	S.D.	Sample II (N=213) Mean	S.D.
1	First Names (M)	9.44	4.507	9.80	4.554
2	Word-Number (M)	4.77	3.602	5.44	3.626
3	Sentences (V)	13.42	4.730	13.75	4.651
4	Vocabulary (V)	27.03	10.317	26.71	10.797
5	Completion (V)	31.97	10.795	31.89	10.581
6	First Letters (W)	36.65	9.778	36.18	11.152
7	Four Letter Words (W)	11.08	4.655	10.85	5.312
8	Suffixes (W)	9.07	4.106	8.46	4.513
9	Flags (S)	25.08	12.427	24.44	11.256
10	Figures (S)	22.70	12.798	22.01	11.451
11	Cards (S)	26.45	13.215	24.85	11.523
12	Addition (N)	16.39	6.991	15.92	7.079
13	Multiplication (N)	32.26	13.430	33.32	12.501
14	Three-Higher (N)	27.21	8.740	25.93	9.840
15	Letter Series (R)	12.40	5.725	12.46	5.718
16	Pedigrees (R)	16.10	7.678	16.45	7.651
17	Letter Grouping (R)	13.32	4.171	13.35	3.879

value is designated by the letter in parentheses. The time limits and scoring formulas are given in Thurstone and Thurstone ([55], p. 28). Product moment correlations among the 17 variables were computed separately for the two samples as shown in Table 2.

Results and Discussion

The first question of interest has to do with the comparability of the means, variances, and covariances for these 17 variables in the two samples. The hypothesis of equal variance-covariance matrices was tested by the procedures given by Anderson ([1], ch. 10) and by Federer [22] and reviewed by Maxwell [41]. The determinant test indicated that the hypothesis of equal variance-covariance matrices was tenable ($\phi = -2 \ln \lambda = 148.697$ for 153 d.f., $p > .05$). The equality of the two sets of means for the 17 variables was evaluated by Hotelling's T^2 statistic ([1], ch. 5). The hypothesis of equal means for the two samples (but not the equality of the means within a sample) was tenable ($F = 1.129$ for 17 and 407 d.f., $p > .05$). Together these two tests indicated that the hypothesis of independent random sampling

TABLE 2

Product Moment Intercorrelations*

Code No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1		472	290	401	299	234	254	296	86	61	52	246	274	250	332	313	297
2	482		189	220	232	209	246	193	44	78	157	151	146	60	238	213	170
3	299	275		833	761	402	275	374	103	19	77	332	297	352	536	567	468
4	331	303	828		772	446	358	473	109	45	105	335	352	384	507	514	404
5	266	273	776	779		394	275	426	342	227	294	329	254	438	490	512	430
6	335	273	439	493	460		627	516	176	104	95	355	365	354	404	365	375
7	342	199	432	464	425	674		480	161	138	49	354	327	318	330	275	317
8	333	290	447	489	443	590	541		79	7	12	288	284	280	327	323	285
9	124	169	117	121	193	178	223	118		672	606	286	189	379	289	277	287
10	32	85	51	77	180	81	192	7	593		728	164	49	236	160	165	181
11	77	193	151	146	174	158	239	114	651	684		171	32	251	200	208	207
12	151	287	268	312	263	241	180	181	208	109	210		651	517	439	320	399
13	259	258	319	344	291	338	295	234	179	144	195	661		546	435	293	452
14	279	223	359	356	342	290	344	298	362	273	331	536	548		512	442	456
15	307	260	447	432	401	381	402	288	252	203	257	361	379	440		671	622
16	447	293	541	537	534	350	367	320	85	129	151	206	298	438	555		538
17	274	216	380	358	359	424	446	325	270	203	293	311	329	410	598	452	

* The data for sample I (N=212) are shown above the principal diagonal and those for sample II (N=213) below the diagonal. Correlations are multiplied by 1000.

from a single multivariate normal distribution was reasonable. The distributions of scores for the factor M tests, variables 1 and 2, however, were somewhat positively skewed. Incidentally, only one of the pairs of sample variances (considered separately for each variable) is significantly different (for variable 11, $F = 1.32$, $.02 < p < .10$).

Since the hypothesis considered by factor analysis is that the 17 variables within each set are not independent, i.e., that one or more degrees of dependence are indicated [4], the hypothesis of independence (the second question) was investigated using the determinant test as given by Anderson ([1], ch. 9) and Bargmann [4] for both samples. The hypothesis of independence was rejected ($\theta = -m \ln V$; $\theta = 1890.303$ and 1857.811 with 136 d.f. for samples I and II, respectively; $p < .001$). The values of the determinants V of the correlation matrices were 8.6658×10^{-5} and 11.8518×10^{-5} for samples I and II, respectively. These results indicated that a factor analysis is justified for each set of data.

Since, for this investigation, the first two hypotheses were tenable, the hypothesized rank of six was then investigated. The canonical correlation (maximum likelihood) approach of Rao [46] was used as a test on the rank,

TABLE 3

Communality Estimates*

Method:	Multiple R squared (inverse)		Centroid high r (adjusted)		Centroid mult. R (15 cycles)		Centroid unity (20 cycles)		Max. like. mult. R (Rao)		Prin. axes unity		Clusters unity	
Sample:	I	II	I	II	I	II	I	II	I	II	I	II	I	II
Code No.														
1	370	390	454	474	382	789	452	606	396	731	712	759	747	751
2	312	323	499	486	762	328	606	417	681	367	794	780	747	751
3	728	747	841	807	838	809	832	805	835	823	869	861	889	876
4	790	756	871	840	853	828	846	836	859	840	885	867	890	879
5	736	699	775	754	796	767	808	772	775	744	864	846	860	846
6	504	569	647	694	643	727	658	736	604	732	754	783	740	776
7	480	550	615	666	626	635	632	631	696	648	778	765	725	755
8	401	455	469	541	452	518	447	501	429	506	634	698	652	700
9	572	515	665	627	607	591	608	585	639	588	750	734	757	738
10	628	551	753	666	819	631	816	651	762	654	830	772	828	763
11	594	591	710	725	672	738	668	728	698	723	796	793	785	797
12	502	533	638	652	550	655	560	688	566	793	758	826	735	757
13	556	523	686	674	810	658	799	656	789	590	813	773	772	759
14	499	512	545	568	522	544	513	535	529	517	653	641	688	697
15	600	507	712	586	736	680	735	640	731	633	784	739	792	752
16	544	543	638	631	612	561	595	612	646	616	761	752	746	694
17	490	451	550	522	535	556	543	562	534	543	705	715	707	703

* Estimates are multiplied by 1000.

i.e., a test on the number of significant factors. The squares of the multiple correlations of each variable with the remaining 16 variables of each set for each of the two samples were computed and used as initial estimates of the diagonal values (in the University of Illinois Illiac program). These values, recommended as lower bounds to the communalities [31], are shown in the first two columns of Table 3. Incidentally, none of the differences between the 17 pairs of corresponding multiple correlation coefficients is significant by Fisher's z transformation ($\sigma_{z_{1i}-z_{2i}} = .102$; for all pairs, $p > .05$).

The hypothesis of not more than five factors was rejected by the χ^2 test used in the Illiac program ($\chi^2 = 105.740$ for sample I and 117.698 for sample II, with critical χ^2 value of 79.9 for 61 d.f., $p < .05$). The hypothesis of only six significant factors, however, was retained since the value of χ^2 quickly dropped below the critical 5 percent value for χ^2 of 66.1 for 49 degrees of freedom (after 5 cycles, the sample I value of $\chi^2 = 56.033$ and the sample II value of $\chi^2 = 52.939$). Comparable results were obtained for Lawley's approximate test for the number of significant factors as given by Thomson [51]; these computations started from the centroid factor matrix (using adjusted high r valuès as estimated communalities) and used two cycles of Bargmann's procedure for determining factor loadings [4]. For both samples, the hypothesis of only five factors ($\chi^2 = 175.831$ for 61 d.f. for sample I and $\chi^2 = 160.592$ for 61 d.f. for sample II) was rejected ($p < .05$) while the hypothesis of six factors (seventh not significant) was tenable (χ^2 = 62.319 for 49 d.f. for sample I and χ^2 = 61.394 for sample II, $p > .05$). The results of these two methods should agree since both are ways of computing solutions of Lawley's maximum likelihood equations. The second procedure illustrates, however, the usefulness and convenience of the centroid method with Bargmann's procedure for testing a given hypothesized rank.

A test of the first ten latent roots was made using Bartlett's test [5] although the results are only of incidental interest since the model being used here is *not* the principal components model. The first ten latent roots for sample I, as computed on the Illiac from the correlation matrix with unity in the diagonals, are 6.31, 2.25, 1.41, 1.27, 1.11, 0.79, 0.58, 0.49, 0.47, and 0.42. The corresponding ten latent roots for the second sample are 6.33, 2.21, 1.42, 1.14, 1.05, 0.95, 0.62, 0.51, 0.44, and 0.39. All ten roots in each sample are significant at the 5 percent level. Kaiser [37] has suggested using the number of latent roots exceeding unity as the number of factors; here that number is five, not six.

After the test of the hypothesized number of factors, the next question of the series is concerned with obtaining an objective statement of the factor pattern of rank six for the six significant factors. Three different aspects of this question can be distinguished: the estimation of communalities, the computation of the factor structure, and (for all but one pair of factor matrices) a further rotational operation. The first phase deals with estimates of the communalities to be inserted in the diagonals prior to factoring. However, in both Bargmann's and Rao's procedures, the iterated factor loadings and, therefore, the communalities, are estimated simultaneously with the test of the number of factors. Since many acrimonious and conflicting statements about the effects of differences in diagonal estimates on factor results have been made, four other sets of estimates of the communalities were computed. Because procedures for iterating communalities converge so slowly, no attempt was made to carry through the iterations to the 50 to 100 cycles that would probably be necessary to obtain convergence to four digits. However, for a given number of factors in a study satisfying the conditions for the existence of a solution, there will be a unique and determinate set of communalities [4].

A criterion of convergence was set arbitrarily at the relatively gross level of a maximum communality difference of $\pm.01$ between two successive cycles. This criterion was met for both samples after five complete cycles of the Illiac program prepared for Rao's procedure. The resulting values are shown in Table 3 in the columns headed with the abbreviations "Max. like., mult. *R*, (Rao)." Since Dr. Kern Dickman (personal communication) has prepared a rapid program for iterating communalities using the centroid factoring procedures, his program was used for two additional sets of estimates. The first of these started with multiple correlation coefficients and required 15 cycles to reach the criteria of a maximum difference of .01 in communalities. As noted earlier, these multiple correlations, shown in the first two columns of Table 3, are the multiple correlations between each variable and the remaining 16 variables for each sample. The resulting iterated values are shown in Table 3 in the column headed "Centroid, mult. *R*, (15 cycles)." Since interest in starting with an arbitrary value such as unity in the diagonals has been expressed, Dickman's procedure was applied to this situation with the results shown in columns of Table 3 headed "Centroid, unity, (20 cycles)." These three iterative solutions seem to be approaching similar limits. Such estimated diagonals need to be compared, however, with those obtained from the widespread (and often ridiculed) practice of inserting the highest correlation coefficient or the highest residual value in each column in the corresponding diagonal cell when the centroid method is used. The results of one cycle of this successive adjustment procedure for six factors are shown in Table 3 in the column headed "Centroid, high r (adjusted)." The remaining two sets of columns in Table 3, one labeled "Prin. axes, unity" and the other labeled "Clusters, unity" are the sums of the squares of the row values in an orthogonal factor matrix of six columns and the squares of a kind of multiple correlation coefficient, both computed with unity in the diagonals of the correlation matrix by means of the principal axes and multiple-group methods, respectively.

The two analyses involving unit diagonals are *not* factor analyses. By definition the unique variances in the factor analysis model are positive and greater than zero; therefore, the diagonal values (communalities) for a factor analysis are in the range $0 < h_i^2 < 1$ [2, 4, 55]. The rotated principal axes solution with unit diagonals is a rotated principal components solution based on the first six latent vectors corresponding to the first six latent roots listed above. The cluster formulation utilizes a set of explicit objective definitions of six linearly and experimentally independent composite variables ([55], p. 63) as an illustration of one possible solution to the fourth source of confusion about factors discussed previously.

The principal axes method is well known and is unambiguous. The multiple-group cluster method, however, requires precise definitions of the clusters and linear function used. The cluster variables (composites) were defined as the average standard scores on the two or three reference tests for each factor. The variables combined are those hypothesized as associated with the PMA factor as indicated by the letter within the parentheses of Table 1. For example, the score for any individual on factor M is defined as the average of the standard scores obtained by that individual (using the means and variances of the appropriate sample) on the two variables, First Names and Word Number. For all other factors, an average of three standard scores would be used to compute (*not* estimate) the individual's factor score. These definitions are readily applied in the computations of the factor loadings on the normals by the multiple-group method using the sums of correlations procedure [8]. The factor loadings as computed are proportional to beta weights from the linear regression model with unit diagonals [7]. It should be noted that part-whole correlations are involved in the computations. The residuals *within* each cluster, including residual diagonals, sum to zero since these cluster vectors are group centroid vectors of the subsets of reference tests.

With the rank and the diagonal values specified, the second phase of determining the factor pattern for the six significant factors can be accomplished using the resulting covariance matrix (the correlation matrix with communalities in the diagonals). Although, theoretically, any method of factoring should be equally effective in reducing the rank of these matrices, some methods are considered as more appropriate than others as judged on the basis of efficiency or of simplicity of computations. The functions used do differ for the different methods, and these differences in method may lead to differences in the simple structure solutions. The methods used were selected, therefore, to provide data relevant to current discussions of the best method of factoring.

Eight analyses using four methods of factoring were made for each sample. These eight analyses included four complete centroid analyses, two principal axes analyses, one canonical correlation analysis, and one multiple-group analysis. The four applications of the centroid method used, as diagonal values, the one-cycle "adjusted high r" values, the 15-cycle values, the 20-cycle values, and Rao's maximum likelihood values. Both Rao's values and unity were used as diagonal entries in the principal axes analyses. Only Rao's values were used in the canonical correlation analysis. These data provide estimates of the effects upon factor loadings of three methods of factoring using a single set of diagonal values (Rao's) and of several variations in diagonals for a single method of factoring (centroid and principal axes). Only a single multiple-group analysis using diagonal values of unity was made as a demonstration of one of the many possible and simple direct solutions to a factor pattern [30]. The direct maximum likelihood solution of Howe for his Model I case ([36], pp. 82–96) was not considered for this study since invariance over diagonal estimates and method of factoring for a single rotational procedure was of primary concern.

The distributions of the 136 sixth-factor residuals from each of these applications of the four factoring methods are shown in Table 4. The means and standard deviations of the residuals are shown at the bottom of the table. Discrepancies between the distributions of residuals for the two samples are clearly shown with sample II having consistently the larger standard deviation. As one might expect, the standard deviations of the residuals computed from the "adjusted high r" centroid method are somewhat, but not markedly, larger than those obtained using the iterated communality estimates. The distributions of residuals from the principal axes factoring method using Rao's maximum likelihood estimates have the smallest variance while the mean values are closest to zero for Rao's factoring procedure and for the centroid factoring method using the 15-cycle diagonal estimates. However, from these distributions of residuals, little preference for one method of factoring over another can be justified, even for the different sets of estimated communalities (excluding unit diagonals).

Since the approximate tests of the number of significant factors given by Burt [15], Cureton [19], and Thomson [51] are especially useful as guides in the searching, subjective, variable-defining, and hypothesis-generating process of an exploratory factor analysis, several of these tests were applied to the residuals and factor loadings of the fifth, sixth, and seventh factors of the original "adjusted high r" analysis of D. C. Bechtoldt. The data for these approximate tests are shown in Table 5. Some question as to the desirability of a seventh factor would be raised by some of these data for sample II in that original analysis. However, McNemar's test of the number of significant factors [43] based upon the ratio of the standard deviation of the distribution of residuals to the average communality agrees with the maximum

TABLE 4

Frequency Distributions, Means, and Standard Deviations of Residuals

	Centroid high r (adjusted)		Centroid • mult. R (15 cycles)		Centroid unity (20 cycles)		Centroid max. like. (Rao)		Max. like. mult. R (Rao)		Prin. axes max. like. (Rao)		Prin. axes unity		Clusters unity	
Sample:	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
Residual																
.09														1		
.08														0		
.07		2		0		0		0		0		0		0		0
.06		2		1		1		1		1		1		2		1
.05		1		1		2		1		1		0	1	1		1
.04	2	2	0	3	1	1	1	3	6	2	1	2	5	3	2	3
.03	7	4	12	3	10	5	9	2	0	6	5	7	10	9	7	5
.02	12	13	13	13	18	14	17	17	11	10	9	9	17	17	7	13
.01	24	25	25	31	16	26	24	31	25	21	32	25	17	19	26	22
00	39	29	28	38	41	38	36	36	53	59	41	45	20	24	35	36
-.01	26	22	32.	25	20	23	20	22	21	17	24	25	18	9	25	15
-.02	14	19	20	7	24	16	22	11	14	7	18	12	13	11	9	12
-.03	8	9	6	9	5	5	7	7	3	8	5	7	11	6	7	6
-.04	2	4	0	3	1	2	0	3	3	2	1	2	4	12	2	5
-.05	2	3		0		3		1		1		1	2	2	1	1
-.06		0		2		0		1		0		0	3	2	1	1
-.07		1		0		0		0		1		0	2	2	2	0
-.08													2	3	1	1
-.09													2	3	2	2
-.10 or less													9	10	9	10
Mean x 1000	-1.4	-1.7	-.02	-0.1	-0.1	-0.3	0.3	0.4	0.1	0.2	-0.7	-0.4	-14.0	-14.2	-13.8	-14.0
S.D. x 1000	17.2	22.7	16.1	18.6	16.5	18.6	16.3	18.5	15.3	17.8	14.5	16.8	44.5	45.2	42.4	45.2

TABLE 5

Data for Approximate Tests of Number of "Significant" Factors

Sample	Factor being tested	From previous factor: Mean residual (less diag.) after sign change	From previous factor: Largest absolute residual	Centroid factor loadings: Largest (absolute)	Centroid factor loadings: Product two largest (absolute)	No. of loadings exceeding critical value* 1.5σ	2σ	3σ
I	5	.028	.190	.318	.087	13	10	7
	6	.022	.109	.338	.086	10	8	2
	7	.007	.049	.186	.031	5	2	0
II	5	.041	.160	.398	.145	14	12	7
	6	.016	.126	.412	.117	8	5	4
	7	.011	.071	.268	.063	6	3	1

* Critical values based on Cureton's solution (19) of Burt's formula for σ_a. The critical values for 1.5σ are .116, .120, and .126 for factors 5, 6, and 7 respectively; for 2σ, the values are .153, .159, and .166; and for 3σ, the values are .223, .232, and .241.

likelihood procedures as to the number of significant factors; the seventh factor would not be significant by his test for any of several analyses using values other than unity in the diagonals. Since in an exploratory study, no proper test of significance of the sequential successive trial type is available, one or two additional factors might indeed be computed as suggested by Thurstone [55] and Rao [46]. Clear-cut residual planes would then aid the investigator in formulating hypotheses for a further confirmatory study using statistical tests of the hypothesized rank [4, 36].

Although the distributions of residuals were very similar from one method of factoring to another within each sample (excepting those methods using unity in the diagonals), the communality estimates shown in Table 3 did differ, especially for the two tests of factor M (rote memory factor). The discrepancies in the communality estimates for variables 1 and 2 call attention to basic and oft-repeated design requirements of factor analysis. These requirements are the necessary and sufficient conditions for identification, i.e., for a unique solution, for the case of one or more common factors as given by Anderson and Rubin [2]. Three or more variables (with nonzero elements) must be used to define a single factor in a factor analysis. (This is not a requirement, however, for the definition of factors by specified linear functions of observed variables as illustrated by the cluster solution since communality estimates are not involved.) For the case of two or more factors, three tests on each factor of a cluster configuration will satisfy the requirements. In the case of factor M, however, there are only two values which both by hypothesis and by the empirical data consistently exceed the definition used here of a zero or near zero factor loading as the range

$\pm .10$. It appears likely that this failure is responsible for the consistent large shifts in the communality estimates for variables 1 and 2 over the two samples for the three iterated sets of estimates, although smaller shifts in the corresponding estimates of other variables did occur.

Given a factor matrix from each of the several factoring operations on each of the two samples, the next phase of determining the factor pattern is to define objectively the oblique simple structure solution. Since the study was designed to provide a cluster configuration, the characteristics of the oblimax solution of Carroll [16] as modified by Pinzka and Saunders [45] should be adequate. The results of the oblimax solution expressed as factor loadings, i.e., as orthogonal projections on normals to the fitted hyperplanes, are presented for three representative factors M, V, and S in Tables 6 to 8, respectively. The three selected tables illustrate the range of variation in sampling fluctuations found in these analyses. These solutions may be termed objective ones since no change was made in any of the oblimax results of the Illiac or IBM 650 output except to define the positive direction of each normal as toward the variables with the highest factor loadings.

The oblimax solutions using only six factors can be compared on these three factors with the graphic solution shown in Table 9 and with the clusters solution given in Tables 6 to 8 in considering the fifth question of interest. The general agreement of these several solutions is clear from an inspection of the data of these tables. With an isolated configuration, the hypothesized

TABLE 6

Rote Memory (M) Factor Loadings*

Method:	Centroid high r (adjusted)		Centroid mult. R (15 cycles)		Centroid unity (20 cycles)		Centroid max. like. (Rao)		Max. like. mult. R (Rao)		Prin. axes max. like. (Rao)		Prin. axes unity		Clusters unity	
Sample: Code No.	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
1	508	463	424	757	501	640	452	717	449	713	447	714	716	708	752	730
2	638	566	820	402	715	475	766	436	763	433	765	439	838	786	808	773
3	-043	-001	-056	-048	-061	-045	-068	-051	-050	-046	-055	-049	-048	-037	-042	-005
4	054	024	025	-013	028	-011	024	-012	007	-012	022	-013	039	000	041	023
5	-019	-022	-026	-048	-020	-046	-021	-046	-004	-056	-008	-052	-009	-044	001	-018
6	-041	005	-058	-003	-057	-004	-050	002	-034	-007	-042	-001	-056	005	-044	-004
7	031	-098	031	-034	035	-021	022	-032	029	-002	029	-019	049	-057	031	-045
8	037	099	036	086	051	113	047	094	020	078	027	088	023	103	013	049
9	-103	101	-067	020	-065	033	-062	022	-074	029	-071	028	-074	056	-052	032
10	027	-078	035	-035	025	-032	026	-035	015	-030	023	-037	029	-031	009	-040
11	105	060	068	010	068	009	070	011	094	-012	085	002	079	035	043	008
12	031	044	-016	-038	-014	-034	-015	-042	005	-064	003	-050	021	012	020	005
13	033	-021	004	-001	005	014	003	013	022	019	021	011	037	026	040	017
14	-120	-049	-094	-002	-083	009	-089	002	114	035	-108	018	-119	-002	-060	-023
15	025	032	031	-011	044	-020	036	-013	034	-027	026	-017	042	-031	017	-032
16	008	079	017	168	024	200	027	171	013	182	010	175	018	162	002	103
17	017	084	004	-051	005	-064	001	-056	-008	-058	-006	-052	-002	-103	-019	-071

* Loadings multiplied by 1000.

TABLE 7

Verbal Facility (V) Factor Loadings*

Method:	Centroid high r (adjusted)		Centroid mult. R (15 cycles)		Centroid unity (20 cycles)		Centroid max. like. (Rao)		Max. like. mult. R (Rao)		Prin. axes max. like. (Rao)		Prin. axes unity		Clusters unity	
Sample: / Code No.	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
1	021	-078	088	-080	083	-086	092	-082	072	-089	064	-092	072	-115	033	-034
2	-010	044	-041	030	-056	005	-051	019	-042	039	-033	040	-055	016	-033	035
3	624	591	580	634	562	626	566	635	600	643	593	640	675	705	694	701
4	649	601	629	625	616	630	630	629	639	633	639	633	696	685	693	680
5	590	571	610	632	625	635	608	618	586	610	589	614	696	719	687	697
6	010	-003	004	-010	006	008	025	-004	024	-005	025	-001	006	018	008	-012
7	-101	-051	-090	-007	-086	-001	-096	-005	-093	-019	-097	-019	-138	-018	-010	-038
8	140	109	152	104	163	102	167	111	151	116	154	116	218	123	092	050
9	-022	-023	-005	-036	007	-044	001	-035	-030	-018	-022	-020	-014	-025	-013	-008
10	-030	004	-024	042	-027	045	-031	042	-020	012	-025	020	-032	039	-028	009
11	035	002	046	007	056	010	051	005	046	-018	045	013	061	015	041	-001
12	038	032	022	-011	017	-003	015	-009	016	020	022	016	032	022	003	013
13	-028	-004	-035	-009	-036	-002	-038	-002	-014	-011	-021	-010	-051	-013	-047	-008
14	040	-001	046	042	051	043	053	040	034	008	040	011	068	017	045	-005
15	-028	-107	-046	-026	-045	-024	-043	-026	-043	-020	-048	-024	004	-003	-004	-037
16	066	103	069	206	075	179	064	202	052	179	051	185	105	257	068	154
17	-042	100	-044	-094	-055	-103	-052	-101	-035	-092	-036	-092	-091	-135	-064	-118

* Loadings are multiplied by 1000.

TABLE 8

Space (S) Factor Loadings*

Method:	Centroid high r (adjusted)		Centroid mult. R (15 cycles)		Centroid unity (20 cycles)		Centroid max. like. (Rao)		Max. like. mult. R (Rao)		Prin. axes max. like. (Rao)		Prin. axes unity		Clusters unity	
Sample: / Code No.	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
1	-019	-051	-039	-018	-030	-028	-035	-024	-032	-034	-036	-029	-042	-065	-042	-052
2	081	122	063	082	063	076	060	084	061	076	064	078	082	097	043	052
3	-105	-042	-124	-032	-124	-034	-118	-030	-102	-025	-106	-028	-111	-028	-104	-035
4	-057	-032	-064	-020	-072	-020	-067	-017	-053	-018	-058	-018	-069	-024	-068	-031
5	186	063	196	081	189	087	194	082	192	075	194	080	199	089	171	066
6	006	-027	008	-039	010	-044	005	-043	002	-043	004	-040	011	-036	022	-019
7	027	036	031	061	032	061	028	057	028	062	027	056	034	070	035	077
8	-056	-031	-074	-052	-073	-055	-075	-054	-064	-048	-068	-048	-065	-057	-058	-058
9	681	708	662	676	664	664	679	674	677	679	677	679	762	785	782	796
10	822	713	873	740	872	757	849	750	835	753	838	749	885	838	885	838
11	775	765	769	775	768	770	781	767	781	765	780	766	854	823	846	828
12	038	-006	048	-038	042	-049	044	-047	051	-046	052	-044	043	-041	021	-045
13	-112	-060	-128	-057	-129	-054	-125	-047	-110	-046	-116	-047	-135	-074	-123	-064
14	096	113	118	133	118	150	122	137	126	131	126	131	133	137	102	109
15	-066	-014	-044	-016	-037	-020	-034	-011	-050	-014	-046	-011	-027	005	-018	022
16	-022	-132	013	-069	014	-067	010	-076	-013	-066	-007	-070	006	-078	010	-079
17	-009	171	016	033	017	028	019	032	004	033	008	033	-008	039	008	057

* Loadings multiplied by 1000.

simple structure for either sample is reproduced with minor variations by any of these techniques. The results of the graphic and cluster methods are consonant with those of the other methods. It should be noted, however, that the graphic solution shown in Table 9 was made in a six-factor subspace of an eight-factor (centroid) structure. Two more factors than hypothesized were computed to compensate for the inefficiency of the centroid method. The six-factor subspace was then set orthogonal to the two thinnest residual hyperplanes defined by the two principal axes corresponding to the two smallest latent roots.

The invariance of the simple structure solution over two samples of cases can be demonstrated by graphical methods, as illustrated for four representative solutions in Figures 1 and 2. Each graph contains 102 points

TABLE 9

Simple Structure Solution by Graphic Techniques

A. Factor Loadings (x 100)

Factor	M		V		W		S		N		R		Estimated Communality 8 factors	
Sample: / Code No.	I	II	I	II	I	II	I	II	I	II	I	II	I	II
1	56	55	-01	-04	-01	08	01	-08	02	-01	06	09	499	552
2	59	55	00	03	02	-03	00	09	-02	02	-03	-06	517	504
3	-07	-01	61	65	-05	-01	-09	-06	-02	-02	08	04	861	818
4	08	02	60	64	04	04	-01	-03	00	02	-06	-04	892	843
5	04	-03	44	62	03	02	10	03	01	-04	-03	01	801	760
6	-06	-06	02	-03	61	61	-01	-08	00	00	04	05	679	702
7	02	-08	-06	-06	62	61	06	08	01	-01	01	10	636	682
8	10	07	05	06	43	48	-07	-10	-01	-01	-02	-07	481	562
9	-06	-01	-09	-04	04	04	61	55	07	01	05	00	679	652
10	00	-05	01	02	05	-02	78	74	-08	00	-05	02	765	698
11	06	03	02	00	-07	02	63	75	-02	00	-01	00	733	740
12	01	-02	02	01	00	-06	00	-01	60	63	-04	-05	648	697
13	03	01	-02	-03	00	05	-07	-02	62	65	01	-05	691	689
14	-05	-02	-05	03	01	01	09	02	37	45	17	08	570	617
15	03	03	-04	05	-01	03	-05	04	07	01	51	52	716	638
16	04	22	02	25	00	-06	-02	-09	-08	00	49	35	647	658
17	00	-04	02	-05	00	17	07	05	10	-02	40	49	564	577

B. Correlations Between Primaries (x 100)*

Factor	M	V	W	S	N	R
M		44	41	07	33	41
V	40		49	06	46	63
W	48	61		10	52	49
S	09	13	13		28	26
N	37	42	38	27		60
R	35	47	42	25	56	

* Sample I values above the principal diagonal and sample II values below.

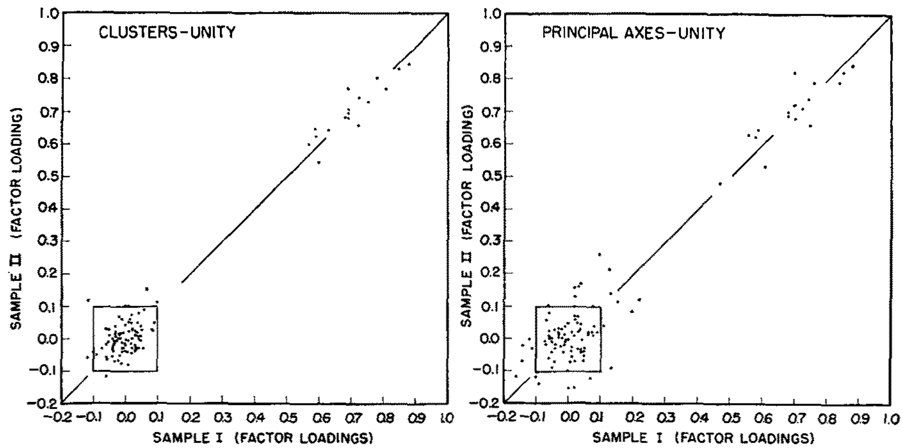

FIGURE 1

Invariance of Factor Loadings for Representative Simple Structure Solutions

(6 factors × 17 variables). The hypothesis of invariance implies that a graph of all factor loadings for one sample plotted against all factor loadings for the second sample for any one solution should show a bivariate distribution with the plotted points symmetrically placed and close to a radial line of 45 degrees. Configurational or even possibly metric invariance, as discussed by Thurstone [55], of the simple structure solution over two samples from the same population is clearly suggested by such graphical techniques for the following four solutions: the clusters with unit diagonals, the principal axes with unit diagonals, the centroid high r (adjusted), and the graphic

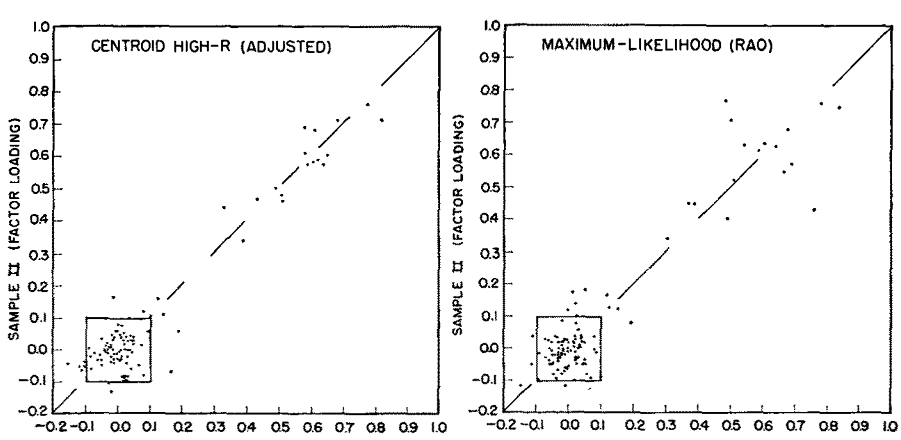

FIGURE 2

Invariance of Factor Loadings for Representative Simple Structure Solutions (judgmental). Somewhat larger deviations for several variables from the 45-degree line will be found for the other five solutions using iterated communalities (e.g., the maximum likelihood plot in Fig. 2). The shifts for the two variables defining factor M are most conspicuous.

Figures 1 and 2 also indicate the number of nonzero loadings in each set of 102 values. Bargmann's [3] value of ±.10 was used to define the zero range drawn on each graph. The clusters solution has the smallest number, 19, of nonzero values (for sample I), while the principal axes, also with unit diagonals, has the largest number, 30 (for sample II). In the graphic solution, sample I has 18 nonzero values (see Table 9). Since the hypothesized value for any one sample was 17, none of the several solutions meets this level when the data for only one sample are considered. However, two solutions, i.e., the clusters solution and the graphic solution, have only 17 *pairs* of loadings *both* greater than ±.10, and the centroid high r solution has only 19 pairs with such loadings. For these three solutions, the agreement between prediction and observation is encouraging with respect both to the number of nonzero values and to the closeness of fit of the data to the 45-degree line. All of the other solutions have more than 19 pairs of loadings greater than ±.10. Under certain conditions, Thurstone's concept of invariance of a simple structure solution [55] receives strong support.

Within a *single* sample, the variation in factor loadings associated with four different sets of communalities and a single factoring method (centroid) is markedly greater than is the variation associated with three methods of factoring using a single set of communalities (Rao's). These results indicate that, for a reasonably well-designed study, the centroid method is not as vastly inferior as has been suggested [37]. The effect upon the factor loadings of variation in diagonal values arising from the use of inaccurate communality estimates, however, is not the essentially irrelevant problem discussed by Wrigley [67] and Guttman [33]. They attempt to factor any arbitrary symmetric matrix and to apply to such a matrix the population rank and communality notions of factor analysis. The communality problem is a pseudo-problem unless the necessary and sufficient conditions are met for the existence of a permissible solution ([4], p. 59) to the factor analysis equations. With empirical data, the question of rank in the population is given a statistical answer under conditions for the existence of a solution.

Unfortunately, the application of available sampling formulations for the evaluation of these variations in an oblique simple structure, within a sample or between samples, either is not appropriate, or as noted by Anderson and Rubin [2], is not feasible at this time. The effect of variation in factor loadings attributable to differences in estimates of communalities (within one sample) is not a proper statistical problem since these variations represent failures to carry the iterations to convergence. Even the data from the clusters solution, however, expressed either as beta weights or, as here, as factor loadings, cannot be evaluated in a within-sample comparison by tests of regression parameters since the factors (the independent variables) are defined by the observed test variables (the dependent variables). Such procedures as developed by Gulliksen and Wilks [28], for example, are appropriate for between-sample comparisons of beta weights or oblique projections when the separation of the independent variables and the dependent variables is maintained in the analysis. The tests of regression parameters in a single-group study are also well known for this case.

Such congruence indices as suggested by Tucker [62] and others are of little value for the matching of factors from the two samples in the current study since the congruence is so uniformly high. Instead, as descriptive statistics of congruence, the second moments (mean squares) of the differences between the pairs of corresponding columns of the rotated factor matrices as well as the second moments of the respective columns were computed. These values are shown in Table 10 for the four representative solutions exhibited in Figures 1 and 2. Tucker's index is defined as a ratio of the sum of the cross products of two columns of factor loadings to the geometric mean of the product of the sums of squares of these same two sets of values; the index, therefore, has been termed an unadjusted correlation coefficient. If desired, such congruence indices can be readily computed from the mean squares of Table 10 by means of the well-known difference formula for a correlation coefficient (without corrections for means). One of the lowest of such congruence indices, that for factor M from the maximum likelihood solution, is .826; the total congruence indices computed over 102 pairs of differences for each of these four representative solutions (in the order given in Table 10) are .987, .967, .945, and .967. Values of this order of magnitude are considered as very acceptable [62].

For any one factor, the mean squares for the sample values tend to be larger for the two solutions using unit diagonals. The larger loadings for the defining variables in these two solutions can be seen also in the figures. In addition, the mean squares of the differences are smallest for the cluster solution. These values reflect the closeness of fit of the points to the 45-degree line. For four factors, the mean squares of the differences for the centroid adjusted solution are next to the smallest although the mean squares for columns also tend to be relatively small. The largest mean squares for columns are found for factor S; this factor also has relatively small mean squares of the differences in the principal axes solution and in the maximum likelihood solution.

Comparable analyses for the several sets of data of this study indicate that differences in the stability of factor loadings do result both from the method of factoring and from the diagonal values used as communality estimates as well as from the characteristics of the data. The effects of the iteration procedures are especially evident in the mean squares of the dif-

TABLE 10

Second Moments (MS)* of Oblique Factor Loadings and of Differences Between Factor Loadings

Method of analysis	Clusters unity			Prin. axes unity			Max. like. mult. R (Rao)			Centroid high r (adjusted)		
Data / Factor	MS_I	MS_{II}	MS_D	MS_I	MS_{II}	MS_D	MS_I	MS_{II}	MS_D	MS_I	MS_{II}	MS_D
M	7.26	6.76	2.39	7.44	7.18	5.86	4.78	4.39	16.02	4.22	3.50	6.17
V	8.55	8.75	1.83	9.12	9.52	5.69	6.86	7.33	3.83	7.02	6.41	3.08
W	8.16	7.88	2.02	7.99	8.16	7.13	5.34	5.57	5.33	5.37	5.61	2.76
S	12.84	12.26	2.55	12.84	12.08	2.95	10.99	9.72	2.94	10.78	9.85	5.24
N	7.95	8.89	2.31	7.83	8.84	6.88	5.08	6.38	9.37	5.06	6.71	3.80
R	6.35	6.59	1.93	6.60	7.10	6.12	3.99	4.08	3.25	4.12	3.87	3.17

* MS_I AND MS_{II} multiplied by 100, MS_D by 1000.

TABLE 11

Congruence Indices for Arbitrary Orthogonal Factor Loadings

Method of analysis / Factors	Prin. axes unity	Max. like. mult. R (Rao)	Centroid high r (adjusted)
I	.995	.995	.996
II	.958	.942	.975
III	.662	.323	.258
IV	.824	.636	.050
V	.686	.859	.933
VI	.846	.800	.683

ferences for factors M and N. The small mean squares of the differences for the cluster solution are suggested as useful reference indices of the uncontaminated sampling fluctuations while the larger mean squares for the other three solutions include the effects of factoring method and of communality estimates. From these data as well as from similar calculations on the other solutions, one might suggest the iterated solutions fit each set of data perhaps too well from the invariance point of view, but such iterated solutions are indicated for the model under consideration.

The results of applying both the graphic and the congruence techniques to the original orthogonal factor matrices indicate little invariance for such values. The results of three such congruence analyses of representative orthogonal factor matrices are illustrated in Table 11. Tucker's index was computed for pairs of columns of the orthogonal factor matrices ordered in terms of decreasing variance contributions of each factor. The inefficiency of the centroid method required reordering of the six "centroid high r" factors as follows: for sample I, factors 1, 2, 4, 3, 5, 6; and for sample II, factors 1, 2, 3, 5, 4, 6. No changes in order of factors were made for the other two sets of calculations, i.e., for the principal axes (unit diagonals) and for the maximum likelihood solutions. The consistency is acceptable only for the first two factors for all three analyses although some consistency is indicated for the other four factors in certain analyses. The poor showing of the centroid "adjusted high r" solution in Table 11 should be contrasted with the very acceptable degree of congruence associated with the mean squares of Table 10 and with the graphs. These data support the frequent suggestion that invariance will not be found for arbitrary orthogonal factor matrices although such invariance may be clearly indicated for a rotated simple structure solution.

The consistency indices above do not differentiate, however, between a simple structure solution and any one of the other possible factor solutions. A more direct attack on the problem of the adequacy of a simple structure solution has been made by Bargmann [3]. He considered the probability of obtaining a given number of vectors within a hyperplane section of small range ($\pm.10$) by rotational methods in a random configuration; the sampling effects (of cases) are not considered. The probability of obtaining a given frequency of zero ($\pm.10$) values of the ratios of factor loadings to length of the vectors (i.e., a_{im}/h_i) in a random configuration was computed by Bargmann for 2 to 12 factors and for 5 to 70 variables, the range of the number of variables varying with the number of factors. For 6 factors and 17 variables (the values for the present study) Bargmann gives the number of (a/h) values in the zero range of $\pm.10$ as 10, 11, and 12 for the rejection at the 5, 1, and 0.1 percent levels respectively of the random configuration hypothesis ([3], p. 18).

The number of ratio values in the critical region for the six factors of the seven oblimax solutions and of the multiple group clusters solution are shown in Table 12. All six factors for both samples would be considered as acceptable by the simple structure 5 percent level criterion for the maximum likelihood and principal axes solutions which use the same communality estimates (i.e., Rao's) and for the multiple-group cluster solution. All but one of the other analyses had only one factor in one of the two samples with only nine ratios in the $\pm.10$ range; the "centroid high r" solution has two unacceptable planes. The number of unacceptable solutions were four for factor S in sample I, one for factor V in sample II, and one for factor R in sample II.

The relatively slight effect of factoring method on the adequacy of the simple structure can be seen in the variation in the number of zero ratios for the three analyses using the same diagonal values (i.e., Rao's). The variations in the number of zero ratios for these three analyses represent differences of ±.03 or less in the values of the ratios. The relatively greater effect upon factor loadings of changes in diagonal values are indicated by the variations among the other analyses.

TABLE 12

Number of (a/h) Ratios in Zero Range*

Method of analysis	Sample	Factor M	V	W	S	N	R
Centroid high (r)	I	12	12	13	9	13	11
	II	12	9	13	10	13	11
Centroid mult. R	I	14	11	14	9	12	12
	II	13	11	12	12	13	13
Centroid Unity	I	14	11	14	9	12	12
	II	13	10	12	12	13	11
Centroid Rao	I	14	11	14	9	12	12
	II	13	11	12	12	13	13
Max. like. Rao	I	13	11	14	10	12	13
	II	13	10	12	12	12	11
Prin. axes Rao	I	13	12	14	10	12	13
	II	13	10	12	12	12	11
Prin. axes Unity	I	14	10	14	10	12	10
	II	13	10	12	12	12	9
Clusters Unity	I	15	13	14	10	13	14
	II	14	12	13	13	14	13
Hypothesized Number		15	14	14	14	14	14

* The number of values in the "zero" range of ±.10 associated with 5%, 1%, and 0.1% "probability" levels are 10, 11, and 12, respectively.

Although three solutions can be considered acceptable simple structure solutions, the data from none of these several analyses agree completely with the hypothesized number of 15 or 14 zero ratios as shown in the last line of Table 12. As noted above, the agreement is better between the number of zero factor loadings (not ratios) and the hypothesized number 15 or 14. A more adequate test of this simple structure hypothesis will probably require the use of such maximum likelihood solutions as are presented by Howe together with further developments of the sampling formulation.

The correlations between the primary axes or factors are shown for each oblimax solution and for the clusters solution in Table 13. The corresponding correlations for the graphic solution are given at the bottom of Table 9. These correlations between the factors are all positive, but the differences from sample I to sample II and from one solution to another within a sample are appreciable. The values do differ somewhat for identical diagonal values (Rao's) as a function of three methods of factoring. However, larger differences

TABLE 13

Correlations Between Pairs of Factors Defined by Primary Axes*

Methods:	Centroid high r (adjusted)		Centroid mult. R (15 cycles)		Centroid unity (20 cycles)		Centroid max. like. (Rao)		Max. like. mult. R (Rao)		Prin. axes max. like. (Rao)		Prin. axes unity		Clusters unity	
Sample:	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
Factor Pairs																
M-V	349	386	334	538	370	567	353	550	329	539	330	535	299	418	342	364
M-W	364	430	347	514	374	527	355	517	326	521	335	518	279	374	334	400
M-S	034	045	048	178	058	191	056	184	048	206	049	190	040	111	105	151
M-N	204	347	221	435	247	441	234	437	196	439	204	423	189	293	259	332
M-R	302	366	296	542	331	580	316	561	281	558	295	544	256	420	353	419
V-W	560	639	580	620	572	599	575	611	575	613	575	613	490	507	493	571
V-S	212	225	201	200	200	191	199	196	206	194	206	191	163	135	179	166
V-N	474	391	499	461	496	431	508	443	496	413	495	417	426	350	436	402
V-R	674	676	714	648	720	646	721	654	710	634	714	634	606	543	617	574
W-S	178	277	174	258	166	260	171	262	184	273	180	275	116	198	124	195
W-N	572	426	583	426	569	404	575	418	583	410	577	418	448	333	462	367
W-R	559	607	558	576	549	568	548	575	567	590	564	595	429	492	465	516
S-N	322	342	310	367	309	362	311	361	313	364	312	360	256	302	262	302
S-R	421	438	356	375	347	389	348	395	387	406	378	404	300	302	289	282
N-R	696	527	683	598	676	573	683	592	704	573	695	577	586	465	573	500

* Correlations multiplied by 1000.

are found for a single method (centroid and principal axes) as a result of changes in communality estimates; these effects on the correlations involving factor M are especially noticeable. The smallest between-sample differences were found for the cluster solutions which reflect most directly the general consistency of the original set of test intercorrelations.

It seems clear that any second- or higher-order analysis will be influenced by the diagonal values used in the first-order analysis (as well as by the number of factors and by the type of preferred solution). Invariance of second-order factor loadings can hardly be expected even from a distinctive isolated configuration unless the first-order factors are explicitly and completely defined as is the case with cluster solution. When the first-order factors are so defined, both the rank and the adequacy of the solutions of the second-order structure can be investigated as in a confirmatory first-order analysis.

The use of orthogonal simple structure [25] or of hierarchical orthogonal solutions [49, 65] does not offer any hope of greater invariance than does an oblique structure since communality estimates are involved in all of these procedures. Analytical solutions for an orthogonal structure are indeed available, but such solutions will exhibit in their first-order factor loadings a combination of the variation found here in oblique factor loadings and in correlations between factors. The forcing of orthogonality between factors in each sample (by definition) also precludes the empirical study of correlations between factors as functions of differences between treatments or populations, differences which Anderson and Rubin [2], Rasch [47], and Thurstone [55] all noted might be associated with changes in these correlations. The regression formulation of factor analysis also indicates the irrelevance of the preference for orthogonal factors. A hypothesis of orthogonality or independence of factors in a population, of course, can be directly evaluated in terms of the correlations between explicitly defined factors in the sample.

The lack of precision of statement in the above discussion of the evaluation of sampling fluctuations is intentional. No sampling formulation for the evaluation of variations in factor loadings and in correlations between factors over both sampling fluctuations and diagonal estimates is currently available. Extensions of the work of Anderson and Rubin, Bargmann, and Howe may lead to more useful sampling formulations in the future. It is suggested that such sampling formulations for existing factor analysis models will require consideration of the several problems developed in this empirical study, i.e., the design of the study, the method for stabilizing communalities, and the method of factoring and of rotation (i.e., the specification of the properties of the preferred solution). However, the restatement of the objective of factor analysis as including the explicit definition of factors changes drastically many sampling problems. Those problems dealing with objectively defined factors are simply the usual univariate or multivariate ones. For other factor theory questions, areas of statistical theory currently under development are relevant. These areas include the identification of parameters of a structure [38] and the fitting of straight lines when both variables are subject to error [40]. When the factors are explicitly defined, these newer analytical developments also become relevant to statements about factors.

The concept of simple structure, however, warrants a brief comment. The theoretical and empirical work of Thurstone and his associates suggests the general usefulness of the concept of simple structure for the variable-defining goal of an exploratory analysis. The objective application of the concept in the current study and the results thereof indicate the possible usefulness of the concept for a confirmatory analysis. The maximum likelihood solutions using good estimators (i.e., unbiased, efficient, etc.) developed by Howe [36] make the concept a precise one. Desired analytical sampling formulations have been indicated and may eventually be developed in a usable form. For these reasons, the rejection by Maxwell [41] of the simple structure concept as not "a precise concept in a valid and efficient statistical theory of factor analysis" seems unduly severe. Under specified and attainable conditions in properly designed confirmatory factor analysis studies with zeros in designated locations, the simple structure concept of factor analysis is indeed offered as a precise concept in an incomplete but valid statistical theory.

The opinion held by Maxwell, however, can be accepted for the vast majority of investigations entitled factor analyses and claiming to use the simple structure concept. These studies, by and large, are exploratory factor analyses (often poorly conceived) for which no statistical tests are available. Variations between investigators in the adequacy of the design of the study, in the procedures for estimating communalities, in the criteria as to when to stop factoring, and in the criteria for rotation, all create differences in the results of the factor analyses. The outcomes of these studies can be represented, at best, by lists of possible reference variables for defining an ever increasing list of factors.

But the list of possible factors is endless, or at least practically so, as emphasized by Thurstone ([54], pp. 194, 201–204, 209; [55], pp. 55–59, 62) and others, since any source of systematic differences between individuals may appear as a factor. A few of these factors, however, may indeed be selected as a stable and useful reference set of concepts accounting for most of the variance of a larger number of variables not used in the definitions of these concepts. The definition of these observable concepts by factor techniques insures some degree of linear independence among them. The usefulness of a proposed set requires in addition, however, evidence of lawful relations derived from experimental laboratory (nonfactorial) investigations of the kind recommended by Thurstone and conducted, for example, to a degree by Stukát [50]. Starting from the available suggested definitions in, say, the ability domain [23], any investigator can provide empirical evidence as to the usefulness of these proposed definitions and of hypotheses involving them.

REFERENCES

[1] Anderson, T. W. *Introduction to multivariate statistical analysis.* New York: Wiley, 1958.

[2] Anderson, T. W. and Rubin, H. Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on mathematical statistics and probability.* Vol. 5. Berkeley: Univ. California Press, 1956. Pp. 111–150.

[3] Bargmann, R. Signifikanzuntersuchungen der einfachen struktur in der faktorenanalyse. *Sonderdruck,* Mit. d. Math. Stat., Physica-Verlag, Wurzburg, 1954.

[4] Bargmann, R. A study of independence and dependence in multivariate normal analysis. Univ. North Carolina, Inst. Statist. Mimeo. Ser. No. 186, 1957.

[5] Bartlett, M. S. Tests of significance in factor analysis. *Brit. J. Psychol., Statist. Sec.,* 1950, **3**, 77–85.

[6] Bechtoldt, H. P. Factor analysis of the Airman Classification Battery with civilian reference tests. *HRRC res. Bull.,* 1953, 53–59.

[7] Bechtoldt, H. P. Statistical tests of hypotheses in confirmatory factor analysis. *Amer. Psychologist,* 1958, **13**, 380. (Abstract)

[8] Bechtoldt, H. P. Construct validity: a critique. *Amer. Psychologist,* 1959, **14**, 619–629.

[9] Bechtoldt, H. P. Comments on "Intraclass correlation vs. factor analytic techniques for determining groups of profiles." *Psychol.Bull.,* 1960, **57**, 157–162.

[10] Bechtoldt, H. P. and Moren, R. I. Correlational analyses of test-retest data with a thirty-year intertrial interval. *Proc. Iowa Acad. Sci.,* 1957, **64**, 514–519.

[11] Bergmann, G. The logic of psychological concepts. *Phil. Sci.,* 1951, **18**, 93–110.

[12] Bergmann, G. *Philosophy of science.* Madison: Univ. Wisconsin Press, 1957.

[13] Brodbeck, M. The philosophy of science and educational research. *Rev. educ. Res.*, 1957, **27**, 427–440.
[14] Brodbeck, M. Models, meaning, and theories. In L. Gross (Ed.), *Symposium on sociological theory*. Evanston: Row Peterson, 1958.
[15] Burt, C. L. Tests of significance in factor analysis. *Brit. J. Psychol., Statist. Sec.*, 1952, **5**, 109–133.
[16] Carroll, J. B. An analytic solution for approximating simple structure in factor analysis. *Psychometrika*, 1953, **18**, 23–38.
[17] Cattell, R. B. Extracting the correct number of factors in factor analysis. *Educ. psychol. Measmt*, 1958, **18**, 791–840.
[18] Cureton, E. E. The principal compulsions of factor-analysts. *Harvard educ. Rev.*, 1939, **9**, 287–295.
[19] Cureton, E. E. A note on the use of Burt's formula for estimating factor significance. *Brit. J. statist. Psychol.*, 1955, **8**, 28.
[20] Dahlstrom, W. G. Research in clinical psychology: factor analytic contributions. *J. clin. Psychol.*, 1957, **13**, 211–220.
[21] Eysenck, H. J. The logical basis of factor analysis. *Amer. Psychologist*, 1953, **8**, 105–114.
[22] Federer, W. T. Testing proportionality of covariance matrices. *Ann. math. Statist.*, 1951, **22**, 102–106.
[23] French, J. W. The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monogr. No. 5*. Chicago: Univ. Chicago Press, 1951.
[24] French, J. W. Manual for kit of selected tests for reference aptitude and achievement factors. Princeton: Educ. Testing Serv., 1954.
[25] Guilford, J. P. *Psychometric methods*. (2nd ed.) New York: McGraw-Hill, 1954.
[26] Guilford, J. P. Three faces of intellect. *Amer. Psychologist*, 1959, **14**, 469–479.
[27] Gulliksen, H. *Theory of mental tests*. New York: Wiley, 1950.
[28] Gulliksen, H. and Wilks, S. S. Regression tests for several samples. *Psychometrika*, 1950, **15**, 91–114.
[29] Guttman, L. General theory and methods for matrix factoring. *Psychometrika*, 1944, **9**, 1–16.
[30] Guttman, L. Multiple group methods for common factor analysis: their basis, computation, and interpretation. *Psychometrika*, 1952, **17**, 209–222.
[31] Guttman, L. "Best possible" systematic estimates of communalities. *Psychometrika*, 1956, **21**, 273–285.
[32] Guttman, L. What lies ahead for factor analysis. *Educ. psychol. Measmt*, 1958, **18**, 497–515.
[33] Guttman, L. To what extent can communalities reduce rank? *Psychometrika*, 1958, **23**, 297–308.
[34] Henrysson, S. *Applicability of factor analysis in the behavioral sciences*. Stockholm: Almqvist and Wiksell, 1957.
[35] Holzinger, K. J. and Harman, H. *Factor analysis*. Chicago: Univ. Chicago Press, 1941.
[36] Howe, W. G. Some contributions to factor analysis. Oak Ridge National Laboratory. ORNL No. 1919, 1955.
[37] Kaiser, H. F. The application of electronic computers to factor analysis. *Educ. psychol. Measmt*, 1960, **20**, 141–151.
[38] Koopmans, T. C. and Reiersöl, O. The identification of structural characteristics. *Ann. math. Statist.*, 1950, **21**, 165–181.
[39] Lawley, D. N. The estimation of factor loadings by the method of maximum likelihood. *Proc. roy. Soc. Edinburgh*, 1940, **60**, 64–82.
[40] Madansky, A. The fitting of straight lines when both variables are subject to error. *J. Amer. statist. Ass.*, 1959, **54**, 173–205.
[41] Maxwell, A. E. Statistical methods in factor analysis. *Psychol. Bull.*, 1959, **56**, 228–235.

[42] Michael, W. B. Symposium: The future of factor analysis. An over-view of the symposium. *Educ. psychol. Measmt*, 1958, **18**, 455-461.

[43] McNemar, Q. On the sampling errors of factor loadings. *Psychometrika*, 1941, **6**, 141–152.

[44] Peel, E. A. Factorial analysis as a psychological technique. *Uppsala Symp. on Psychol. Factor Analysis*, 1953, 7–22.

[45] Pinzka, C. and Saunders, D. R. Analytical rotation to simple structure II. Extension to oblique solution. Princeton: Educ. Test Serv., 1954.

[46] Rao, C. R. Estimation and tests of significance in factor analysis. *Psychometrika*, 1955, **20**, 93–111.

[47] Rasch, G. On simultaneous factor analysis in several populations. *Uppsala Symp. on Psychol. Factor Analysis*, 1953, 65–71.

[48] Reiersöl, O. On the identifiability of parameters in Thurstone's multiple-factor analysis. *Psychometrika*, 1950, **15**, 121–149.

[49] Schmid, J. and Leiman, J. M. The development of hierarchical factor solutions. *Psychometrika*, 1957, **22**, 53–61.

[50] Stukát, K. G. *Suggestibility: a factorial and experimental analysis.* (*Acta Psychologica Gothoburgensia*, II) Stockholm: Almqvist and Wiksell, 1958.

[51] Thomson, G. *The factorial analysis of human ability.* (5th ed.) Boston: Houghton Mifflin, 1953.

[52] Thurstone, L. L. *The vectors of mind.* Chicago: Univ. Chicago Press, 1935.

[53] Thurstone, L. L. Current misuse of the factorial methods. *Psychometrika*, 1937, **2**,73–76.

[54] Thurstone, L. L. Current issues in factor analysis. *Psychol. Bull.*, 1940, **37**, 189–236.

[55] Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. Chicago Press, 1947.

[56] Thurstone, L. L. and Thurstone, T. G. Factorial studies of intelligence. *Psychometric Monogr. No. 2.* Chicago: Univ. Chicago Press, 1941.

[57] Thurstone, L. L. and Thurstone, T. G. *Primary mental abilities* (manual and tests). Chicago: Sci. Res. Assoc., 1950.

[58] Tryon, R. C. Communality of a variable: formulation by cluster analysis. *Psychometrika*, 1957, **22**, 241–259.

[59] Tryon, R. C. General dimensions of individual differences: cluster analysis vs. multiple factor analysis. *Educ. psychol. Measmt*, 1958, **18**, 477–495.

[60] Tryon, R. C. Domain sampling formulation of cluster and factor analysis. *Psychometrika*, 1959, **24**, 113–135.

[61] Tucker, L. R. The role of correlated factors in factor analysis. *Psychometrika*, 1940, **5**, 141–152.

[62] Tucker, L. R. A method for synthesis of factor analysis studies. PRS Report No. 984. Dept. Army: AGO, Personnel Res. Sec., Washington, D. C., 1951.

[63] Tucker, L. R. The objective definition of simple structure in linear factor analysis. *Psychometrika*, 1955, **20**, 209–225.

[64] Tucker, L. R. An inter-battery method of factor analysis. *Psychometrika*, 1958, **23**, 111–136.

[65] Wherry, R. J. Hierarchical factor solutions without rotation. *Psychometrika*, 1959, **24**, 45–51.

[66] Wrigley, C. Objectivity in factor analysis. *Educ. psychol. Measmt*, 1958, **18**, 463–476.

[67] Wrigley, C. The effect upon the communalities of changing the estimate of the number of factors. *Brit. J. statist. Psychol.*, 1959, **12**, 35–51.

[68] Young, G. and Householder, A. S. Factorial invariance and significance. *Psychometrika*, 1940, **5**, 47–56.

Manuscript received 6/7/60

Revised manuscript received 3/1/61