

Comparisons of Order Analysis and Factor Analysis in Assessing the Dimensionality of Binary Data

Steven L. Wise
Southern Illinois University

Previous research has not shown a clear relationship between order analytic and factor analytic approaches to assessing the dimensionality of binary data. This study compared factor analysis with three order analysis procedures. Comparisons were based on eight datasets with known dimensionality and two multidimensional sets of mathematics data. Two of the order analysis procedures fared poorly in reproducing the factor structure of the datasets. The third procedure reproduced the factors for datasets with orthogonal factors but failed to reproduce the factors for datasets containing oblique factors. Reasons for the differences between these procedures are discussed.

There has recently been renewed interest in methods for determining the dimensionality of binary test data. The strong impact of item response theory (IRT) on psychological and educational measurement has led to a critical examination of the unidimensionality assumption that underlies the IRT methods currently in use. This assumption is needed because of estimation problems involved with the multidimensional IRT models that are available. Hence, the current state of IRT prevents the measurement of more than one latent trait at a time. However, many tests yield multidimensional data, implying that more than one latent trait would be necessary to account for test performance.

One possible solution to this problem of multidimensionality would be to extract unidimensional subsets from the total set of test items and to use IRT to generate separate ability estimates for each dimension. Although factor analysis remains the most commonly used method for extracting unidimensional sets of items, other methods are available, including order analysis (Krus, 1975). Order analysis makes use of the logical relations between elements of a binary data matrix in constructing unidimensional orders of the elements. These unidimensional orders approximate Guttman scales.

The logical relation of interest in order analysis is the order or dominance relation. In the context of ability testing, if a person correctly answers item i , he or she is said to dominate that item. Likewise, if the person does not correctly answer item i , then he or she is dominated by item i . However, in order analysis interest is in using the observed dominance relations between persons and items to determine unobserved, implied relations between persons and items within the same set. For example, if a person incorrectly answers item i and correctly answers item j , then it is implied that item i dominates item j for that person. If a joint placement of persons and items on a single ability dimension is considered, a more parsimonious conceptualization results. When item i dominates item j , then more ability is needed to correctly answer i than is needed to correctly answer j (this can also be stated as “ i is more difficult than j ”).

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 7, No. 3, Summer 1983, pp. 311–321
© Copyright 1983 Applied Psychological Measurement Inc.
0146-6216/83/030311-11\$1.80

Item dominance relations should be consistent across persons in a sample. A dominance relation between items i and j (termed an ij dominance) is indicated only if there is a sufficient preponderance of persons showing an ij dominance over persons showing a ji dominance. Krus (1977) suggested the use of McNemar's (1947) z statistic for correlated proportions to evaluate the presence of a dominance relation.

Krus (1977) stated that "the dimensionality of a set of data can be derived from asymmetric, transitive, and connected relations between its elements, i.e., from those relations which define an order" (p. 587). There are times when there are not clear dominances between all pairs of items. According to Krus (1975), this indicates that the items are not members of the same order and that the data are multidimensional. Based on this, a deterministic order analytic model for determining the dimensionality of item sets (Krus & Bart, 1974), and later a probabilistic model (Krus, 1977), were developed. Based on the deterministic model, the probabilistic model generates "order loadings" for the items on each dimension. Subsets of persons are isolated, with each subset reflecting a different item order (dimension). For a given dimension, the order loadings reflect the relative order position of each item. However, this procedure isolates subsets of persons rather than subsets of items. In the context of the present paper, the parallel procedure of isolating item subsets (yielding different person orders) is more appropriate. In this case the order loadings would reflect the relative standing of persons on the unidimensional item scales.

Cliff (1977) developed a series of indices for evaluating the consistency of orders across items (or equivalently, persons). Reynolds (1981) outlined an algorithm, using one of Cliff's (1977) indices for extracting item chains. Items are combined into chains such that the individual chain consistency remains above some minimally acceptable level.

Comparisons of Order Analysis and Factor Analysis

There have been a number of studies comparing order analysis with factor analysis. However, al-

most all of the studies involved use of Krus' (1977) probabilistic model with subsets of persons being isolated rather than items. Krus and Tellegen (1975) investigated dimensions of social consciousness and found similar results using order analysis and factor analysis, as did Krus, Weidman, and Bland (1975) for dimensions of institutional evaluation. Krus and Weiss (1976) found congruence between order analysis and factor analysis for Thurstone's (1947) "box data" but found differing results for the two procedures using random data generated by the method reported in Armstrong and Soelberg (1968). Bart (1978) reanalyzed the data reported in Bock and Lieberman (1970) and concluded that the factor structure of a set of data did not appear to relate in any clear way to the order structure. Krus (1978) presented results for a marital adjustment checklist that suggested moderate similarity between order structure and factor structure. Reynolds (1981) found encouraging results for his procedure using Cliff's (1977) c_{11} index for a social distance questionnaire.

The results of the above studies suggest a lack of a clear relationship between order analysis and factor analysis. This study was designed to compare factor analytic results with results obtained using various order analysis procedures in order to delineate any differences that exist among the procedures.

Method

Simulated Datasets

Seven simulated orthogonal datasets, consisting of 10 items and 500 persons, were computed as follows. A factor pattern matrix and a vector of unique factor loadings were prespecified, from which a population variance-covariance matrix was generated using a modified Tucker, Koopman, and Linn (1969) procedure that simulated the effects of random error on the variance-covariance matrix by allowing for the influence of a number of minor random factors. This population variance-covariance matrix was then used in conjunction with a vector of prespecified population item means to generate dichotomous item scores from a multivariate normal population.

The seven simulated datasets are described in Table 1. It was anticipated that the distributions of item difficulty levels might have a differential effect on the order analysis procedures. Hence, three types of item mean distributions were used: widely spaced means where each item difficulty level was very distinct from that of the other items, moderately spaced means where some item difficulty levels were similar, and narrowly spaced means where all of the item difficulty levels were similar. For datasets with widely spaced means, all pairs of item difficulty levels differed by at least .08. For moderately spaced means, some of the pairs of item difficulty levels differed by .05 or less, with the remaining pairs differing by at least .08. For narrowly spaced means, all pairs of item difficulty levels differed by less than .05. Also, for the two-factor datasets (H2, M2, L2), Items 1 through 4 always loaded on one factor, and Items 5 through 10 loaded on the other factor.

Dataset M10 was unique in that it was generated so that there were no common factors among the items. It consisted of 10 unrelated items with moderately spaced means. This dataset was useful in comparing order analysis procedures in their capability of indicating a lack of order structure.

Signed Numbers Mathematics Data

The mathematics data consisted of 16 dichotomous mastery scores derived from a 64-item signed numbers test administered to 125 eighth-grade stu-

dents. There were 16 skills, each measured by four parallel items. Examples of these skills are shown in Table 2. If a student answered at least three of the four items correctly, he or she was deemed a master of that skill and given a mastery score of one. Otherwise, a score of zero was given (non-mastery).

Two forms of the signed numbers dataset were analyzed. Birenbaum and Tatsuoka (1980) have described a procedure for detecting inappropriate strategies used by students in solving signed number problems. Often, students can obtain "correct" answers to some of these problems using incorrect strategies. Once an incorrect strategy was detected for a given student, it was possible to determine the items for which the student would have given the correct answer using the inappropriate strategy. An "adjusted" dataset was then constructed from the original 64-item signed numbers dataset such that items deemed to have been answered correctly by an inappropriate strategy were rescored as incorrect. Order analyses and factor analyses were subsequently performed on both the unadjusted (UMATH) and adjusted (AMATH) 16-item mastery datasets.

Order Analysis Procedures

Three order analysis procedures were used: the deterministic order analysis method of Krus and Bart (1974), Reynolds' (1981) algorithm using c_{11} as an extraction index, and Reynolds' procedure

Table 1
Description of the Simulated Datasets

Dataset Label	Description
H1	One factor with wide spacing between the item means.
M1	One factor with moderate spacing between the item means.
L1	One factor with narrow spacing between the item means.
H2	Two factors with wide spacing between the item means.
M2	Two factors with moderate spacing between the item means.
L2	Two factors with narrow spacing between the item means.
M10	Consisted of ten independent items with moderate spacing between the means (essentially a 10-dimensional dataset).

Table 2
Examples of the 16 Signed-Number
Mathematics Skills

Item (Skill)	Example	Operation
1	$1 - (-10) = 11$	Subtraction
2	$9 - (-7) = 16$	Subtraction
3	$-7 - 9 = -16$	Subtraction
4	$-12 - 3 = -15$	Subtraction
5	$-3 - 12 = -15$	Subtraction
6	$-6 - (-8) = 2$	Subtraction
7	$-16 - (-7) = -9$	Subtraction
8	$8 - 6 = 2$	Subtraction
9	$2 - 11 = -9$	Subtraction
10	$6 + 4 = 10$	Addition
11	$-14 + (-5) = -19$	Addition
12	$-5 + (-7) = -12$	Addition
13	$-3 + 12 = 9$	Addition
14	$-6 + 4 = -2$	Addition
15	$12 + (-3) = 9$	Addition
16	$3 + (-5) = -2$	Addition

using c_{13} . To determine the presence of a relation in Krus and Bart's (1974) procedure, a criterion value of 1.64 was used for McNemar's (1947) z . Krus's (1977) probabilistic order analysis procedure was not used for two reasons. First, it was decided that the results obtained from the deterministic and probabilistic models would be similar enough so that both procedures would not be necessary in this study. Second, since Reynolds' (1981) method is deterministic, the deterministic order analysis method was chosen in order to permit the most straightforward comparisons among the results of the different methods. For the c_{11} and c_{13} procedures, cutoff values of .90 and .70, respectively, were used. These are relatively high cutoff values; however, preliminary simulation analyses indicated that these values resulted in the closest congruence between chains extracted using Reynolds' procedure and factor analysis. Table 3 lists, for each dataset, the overall values of c_{11} and c_{13} , and KR20.

Results

Simulated Data

In order to verify the factor structures of the simulated datasets, factor analyses of the matrices of phi coefficients were performed, using a principal factor solution with iterated communality estimates. Initial communality estimates were the squared multiple correlations of each item with the other items in the dataset. Phi coefficients were used rather than tetrachoric coefficients because of persistent problems with tetrachoric coefficients. When one item dominates another item in a perfectly consistent manner (i.e., no counterdominances), the tetrachoric correlation is equal to 1.0. However, since in most cases the correlation coefficient is calculated for sample data, there would typically be reluctance to accept 1.0 as a population correlation estimate. Also, matrices of sample tetrachoric coefficients will often be non-Gramian,

Table 3
Overall Consistency Statistics for
the Nine Datasets

Dataset	c_{t1}	c_{t3}	KR20
H1	.972	.944	.875
M1	.863	.901	.942
L1	.071	.745	.965
H2	.655	.393	.749
M2	.755	.374	.703
L2	.036	.329	.826
M10	.606	-.011	-.057
UMATH	.678	.509	.866
AMATH	.777	.764	.936

in violation of basic assumptions of the factor analytic model. Neither of these problems occurs when phi coefficients are used. Although phi coefficients are influenced by the relative difficulty levels of the items, Comrey (1973) reported finding the influences of difficulty factors to be minor and endorsed the use of phi rather than tetrachoric coefficients. Hence, phi coefficients were deemed to be appropriate in this study.

For datasets where more than one common factor was extracted, factors were rotated using the Varimax criterion. All seven datasets showed clear factorial dimensionality in agreement with the factor pattern matrices from which the datasets were generated. For Dataset M10, a scree test of the eigenvalues led to the conclusion that no common factors were present.

The order analysis results for Datasets H1, M1, and L1 are shown in Table 4. For H1, all three procedures correctly extracted a single chain (dimension) of items. For M1 the three procedures were not in agreement. While the single chain was correctly extracted using c_{t3} , use of the other two procedures yielded multiple chains. However, if the minimum consistency level of c_{t1} is lowered from .90 to .86, then the correct single chain would have been extracted for the c_{t1} procedure. For Dataset L1, composed of items that were highly similar in terms of difficulty level, the c_{t3} procedure was

the only procedure that extracted the single dimension. The other two procedures failed to determine any item chains. Note in Table 3 that the overall value of c_{t1} was near zero, while c_{t3} was fairly high.

Table 5 shows the order analysis results for Datasets H2, M2, and L2. For H2, Krus and Bart's (1974) procedure could not accurately extract the two factors; Items 6, 7, and 9 were incorrectly combined in a chain with Items 1, 2, 3, and 4. Reynolds' (1981) procedure extracted the correct chains when either c_{t1} or c_{t3} was used. For M2, however, only the c_{t3} procedure extracted the two dimensions. The c_{t1} procedure extracted one of the dimensions but could not extract the other. The results for Krus and Bart's (1974) procedure were chaotic in terms of the factor structure of this dataset. For Dataset L2, as for L1, only the c_{t3} procedure correctly extracted the two dimensions. The other two procedures failed to combine any items into chains.

The chain extraction results for M10, shown in Table 6, illustrated other differences among the three procedures. In this dataset, there were no real common factors present among the items. The c_{t3} procedure extracted no chains at all. Krus and Bart's procedure, however, yielded a large (eight-item) chain, and the c_{t1} procedure yielded a number of small chains.

Table 4
Item Chain Extraction for Datasets H1, M1, and L1

Dataset	Krus & Bart Procedure	c_{t1} Procedure	c_{t3} Procedure
H1	(1-2-3-4-5-6-7-8-9-10)	(1-2-3-4-5-6-7-8-9-10)	(1-2-3-4-5-6-7-8-9-10)
M1	(2-3-4-5-6-7)	(1-2-3-4-5-6-7-8-9-10)	(1-2-3-4-5-6-7-8-9-10)
	(1)		
	(8)		
	(9)		
	(10)		
L1	(1)	(1)	(1-2-3-4-5-6-7-8-9-10)
	(2)	(2)	
	(3)	(3)	
	(4)	(4)	
	(5)	(5)	
	(6)	(6)	
	(7)	(7)	
	(8)	(8)	
	(9)	(9)	
	(10)	(10)	

Note: Cutoff values of c_{t1} and c_{t3} were .90 and .70, respectively.

All of the items loaded on a single factor.

Signed Numbers Data

Factor analyses of the matrices of phi coefficients for the two signed numbers datasets are shown in Tables 7 and 8. For the UMATH dataset, a two-factor solution is presented, although a scree test of the eigenvalues did not clearly suggest the number of factors to extract. Factor solutions were obtained for two through five factors, and the two-factor solution best approximated simple structure. The subtraction items (1 through 9) constituted one factor, while four of the addition items (13 through 16) constituted the second factor. The four second-factor items were all skills dealing with the addition of two numbers that were opposite in sign.

However, when the data were adjusted for presumably "erroneously correct" responses (AMATH), two clear factors of subtraction and addition emerged. Only two eigenvalues were greater than one, and a very clear simple structure was present. The correlation between the two factors was .46.

Order analyses of the signed numbers data, shown in Table 9, gave very different results from those of the factor analyses. For both datasets neither the Krus and Bart (1974) procedure nor the c_{t1} procedure yielded chains that showed any resemblance

to the factors. The c_{t3} procedure also failed to reproduce the factor structure for either dataset. For AMATH in particular, the c_{t3} procedure found one chain with fairly high overall consistency ($c_{t3} = .764$). Various other cutoff values of c_{t3} were used in analyzing the AMATH dataset, none of which yielded chains that reproduced the factor structure of AMATH.

Because it was suspected that the failure of the c_{t3} procedure to reproduce the factors for the mathematics datasets was due to their oblique factor structures, an eighth simulated dataset was constructed. This dataset, labeled OBL1, was composed of two factors containing 15 items each ($N = 500$) with a correlation of .52 between the factors. A factor analysis using an Oblimin rotation confirmed the presence of the two factors. When the OBL1 dataset was analyzed using the c_{t3} procedure, a multitude of item chains was extracted, none of which remotely resembled either of the factors.

Discussion

It quickly became obvious from the results of the simulated data that Krus and Bart's (1974) pro-

Table 5
Item Chain Extraction for Datasets H2, M2, and L2

Dataset	Krus & Bart Procedure	c_{t1} Procedure	c_{t3} Procedure
H2	(1-6-2-7-3-9-4) (5-8-10)	(1-2-3-4) (5-6-7-8-9-10)	(1-2-3-4) (5-6-7-8-9-10)
M2	(6-1-7-4-9) (5-8-10) (2) (3)	(1-3-4-9) (1-2-4-9) (5-6-7-8-9-10)	(1-2-3-4) (5-6-7-8-9-10)
L2	(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)	(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)	(1-2-3-4) (5-6-7-8-9-10)

Note: Cutoff values of c_{t1} and c_{t3} used were .90 and .70, respectively.
Items 1-4 loaded highly on one factor, and items 5-10 loaded highly on the other factor.

cedure did not perform very well in reproducing the factor structures of the datasets. It reproduced the factors only for Dataset H1. The c_{t1} procedure did not fare much better; it reproduced the factors only for Datasets H1 and H2. Basically, there are two reasons for the poor results from these two procedures. First, when a factor contains two or more items with highly similar difficulty levels, not all of these items will frequently appear in the same chain. Items that are too close together in

terms of difficulty will often fail to show a clear dominance relation, indicated by a low value of McNemar's z . Hence, by Krus and Bart's procedure, this absence of a relation will imply that the items do not belong to the same dimension. Correspondingly, the low z value also means that c_{t1} between the items will also be very low. Thus, items similar in difficulty often show very inconsistent dominance relations.

In fact, if two items, i and j , have equal difficulty

Table 6
Item Chain Extraction for Dataset M10

Dataset	Krus & Bart Procedure	c_{t1} Procedure	c_{t3} Procedure
M10	(1-2-3-6-7-8-9-10) (4) (5)	(1-2-10) (1-3) (1-7) (1-8) (1-9) (4) (5) (6)	(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

Note: Cutoff values of c_{t1} and c_{t3} used were .90 and .70, respectively.
This dataset contained no common factors.

Table 7
Principal Factor Analysis of Phi Coefficients
for the Unadjusted Signed-Numbers (UMATH) Dataset

Item Number	Item Mean	Item S.D.	Factor Loadings	
			Factor I	Factor II
1	.648	.480	.817	-.119
2	.680	.468	.847	-.120
3	.584	.495	.696	-.078
4	.576	.496	.698	-.162
5	.720	.451	.891	-.096
6	.744	.438	.713	.082
7	.824	.382	.617	.068
8	.856	.352	.485	-.025
9	.704	.458	.635	.118
10	.992	.089	.119	.037
11	.912	.284	.368	.159
12	.936	.246	.352	.060
13	.920	.272	.038	.765
14	.944	.231	-.011	.591
15	.920	.272	-.035	.509
16	.920	.272	.051	.684

Note: Factors were rotated using the Oblimin method.

levels, there simply cannot be a dominance relation between them, since there must be the same number of ij dominances as ji dominances. It makes no difference whether or not the two items are from the same factor. This explains, in part, why Krus and Bart's procedure and the c_{ij} procedure performed so poorly. H1 and H2 were the only datasets containing items that were all very distinct in terms of difficulty level. Hence, for these datasets the problem of similar difficulty levels is minimal.

This problem of similar difficulty levels can also be reconsidered in a fashion analogous to that of statistical power. As the population means for two treatment groups become more similar, other things being equal, the power to detect the difference between these means decreases. Likewise, as two item population means become more similar, the power to detect the fact that one item is actually

more difficult than (i.e., dominates) the other, decreases.

The second problem with Krus and Bart's procedure and the c_{ij} procedure is also related to the distribution of item difficulty levels. Two items that are statistically independent can show a consistent dominance relation that is due solely to difficulty differences between the two items. This is the same problem that led to consideration of chance reproducibility in scalogram analysis (Chilton, 1969; Schooler, 1968; Schuessler, 1961; White & Saltz, 1957). For example, consider two items that are independent and have difficulty levels of .30 and .90 computed from a sample of 100 persons. The expected number of persons for whom Item 1 dominates Item 2 is equal to $100 \times (\text{probability of incorrectly answering Item 1 and correctly answering Item 2}) = (100)(.70)(.90) = 63$. Likewise,

Table 8
Principal Factor Analysis of Phi Coefficients
for the Adjusted Signed-Numbers (AMATH) Dataset

Item Number	Item Mean	Item S.D.	Factor Loadings	
			Factor I	Factor II
1	.600	.492	.834	.015
2	.624	.486	.865	.007
3	.536	.501	.771	-.022
4	.536	.501	.770	-.024
5	.688	.465	.937	.030
6	.664	.474	.895	.041
7	.696	.462	.920	.050
8	.792	.408	.679	-.068
9	.648	.480	.749	.068
10	.960	.197	.144	.353
11	.888	.317	.108	.745
12	.904	.296	.069	.786
13	.888	.317	-.086	.858
14	.896	.306	-.008	.850
15	.872	.335	-.094	.814
16	.880	.326	-.014	.833

Note: Factors were rotated using the Oblimin method.

the expected number of persons for whom Item 2 dominates Item 1 is equal to $(100)(.30)(.10) = 3$. In this case, $z = 7.39$ and $c_{i1} = .91$. This illustrates that items that are disparate in difficulty will tend to show consistent dominance relations regardless of whether or not they belong to the same factor. This problem was most clearly manifested in Dataset M10, which was composed of 10 independent items.

These two problems undermine the usefulness of both Krus and Bart's (1974) procedure and the c_{i1} procedure for determining item chains that are comparable to factors. Although the two procedures extract chains that approximate Guttman scales, it is clear that Guttman scales are not equivalent to factors.

In the above mentioned example with two independent items ($p_1 = .30$, $p_2 = .90$) the value of c_{i3} is 0. This illustrates a very desirable property

of c_{i3} ; that the expected number of chance dominances (for independent items) is taken into consideration. The c_{i3} procedure is also less prone to the first problem described above: that items too similar in difficulty level tend not to show a clear dominance relation.

The c_{i3} procedure yielded chains that correctly reflected the factor structures for all seven simulated orthogonal datasets. It was found to be consistently superior to both the c_{i1} and Krus and Bart (1974) procedures. The better performance of c_{i3} compared with c_{i1} is in agreement with results found by Cudeck (1980). However, in the case of correlated factors, the c_{i3} procedure may not be able to distinguish between items loading on different factors.

It is evident that order analysis and factor analysis can yield substantially different results regarding dimensionality. The problem of similar dif-

Table 9
Item Chain Extraction for the Mathematics Datasets

Dataset	Krus & Bart Procedure	c_{t1} Procedure	c_{t3} Procedure
UMATH	(4-1-5-7-11-10) (3-2-8-15) (9-13) (6-16) (12) (14)	(4-5-11-12-10) (3-5-11-12-10) (1-2-12-10) (3-5-14-10) (9-12-10) (4-8-10) (1-7-10) (6-10) (13-10) (15)	(4-3-1-2-5-6-11-12-10) (4-1-5-6-7-8-10) (9-11-12) (5-14-10) (13) (15) (16) (12)
AMATH	(3-1-6-8-15-10) (4-2-5-16) (9-11) (7-13) (14) (12)	(3-4-6-5-7-11-12-10) (3-6-5-7-14-12-10) (1-2-5-7-11-12-10) (4-6-5-8-10) (2-7-16-10) (9-14-10) (5-13-10) (15)	(3-4-1-2-9-6-5-7-8-15-16-11-13-14-12-10)

Note: Cutoff values of c_{t1} and c_{t3} used were .90 and .70, respectively.

ficulty levels suggests that proximity relations between items should be considered along with dominance relations in evaluating dimensionality. To this end, Krus and Krus (1980) have described a procedure where items are grouped using a proximity measure and then order analyzed using Krus's (1977) method. Also, Wise (1982) outlined an algorithm, using both proximity and dominance relations, for extracting item chains that form partial orders. Wise found considerable similarity between factor structure and partial order structure. Hence, it appears that proximity relations should not be ignored when evaluating dimensionality.

References

- Armstrong, J. W., & Soelberg, P. On the interpretation of factor analysis. *Psychological Bulletin*, 1968, 70, 361-364.
- Bart, W. M. An empirical inquiry into the relationship between test factor structure and test hierarchical structure. *Applied Psychological Measurement*, 1978, 2, 331-335.
- Birenbaum, M., & Tatsuoka, K. K. *The use of information from wrong responses in measuring students' achievement* (Research Report 80-1). Urbana IL: Computer-based Education Research Laboratory, 1980.
- Bock, R., & Lieberman, M. Fitting a response model of n dichotomously scored items. *Psychometrika*, 1970, 35, 179-197.
- Chilton, R. J. A review and comparison of simple statistical tests for scalogram analysis. *American Sociological Review*, 1969, 34, 238-245.
- Cliff, N. A theory of consistency or ordering generalizable to tailored testing. *Psychometrika*, 1977, 42, 375-399.
- Comrey, A. L. *A first course in factor analysis*. New York: Academic Press, 1973.
- Cudeck, R. A comparative study of indices for internal consistency. *Journal of Educational Measurement*, 1980, 17, 117-130.
- Krus, D. J. *Order analysis of binary data matrices*. Los Angeles: Theta Press, 1975.
- Krus, D. J. Order analysis: An inferential model of dimensional analysis and scaling. *Educational and Psychological Measurement*, 1977, 37, 587-601.
- Krus, D. J. Logical basis of dimensionality. *Applied Psychological Measurement*, 1978, 2, 321-329.
- Krus, D. J., & Bart, W. M. An ordering theoretic method of multidimensional scaling of items. *Educational and Psychological Measurement*, 1974, 34, 525-535.
- Krus, D. J., & Krus, P. H. Dimensionality of hierarchical and proximal data structures. *Applied Psychological Measurement*, 1980, 4, 313-321.
- Krus, D. J., & Tellegen, A. Consciousness III: Fact or fiction? *Psychological Reports*, 1975, 36, 23-30.
- Krus, D. J., Weidman, J. C., & Bland, P. C. SSIE:

- Semiprojective scales of institutional evaluation. *Research in Higher Education*, 1975, 3, 131-138.
- Krus, D. J., & Weiss, D. J. Empirical comparison of factor and order analysis on prestructured and random data. *Multivariate Behavioral Research*, 1976, 11, 95-104.
- McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 1947, 12, 153-157.
- Reynolds, T. ERGO: A new approach to multidimensional item analysis. *Educational and Psychological Measurement*, 1981, 41, 643-660.
- Schooler, C. A note of extreme caution on the use of Guttman scales. *American Journal of Sociology*, 1968, 74, 296-301.
- Schuessler, K. F. A note on the statistical significance of scalogram. *Sociometry*, 1961, 24, 312-318.
- Thurstone, L. L. *Multiple factor analysis*. Chicago: University of Chicago Press, 1947.
- Tucker, L. R., Koopman, R. F., & Linn, R. L. Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, 1969, 34, 421-459.
- White, B. W., & Saltz, E. Measurement of reproducibility. *Psychological Bulletin*, 1957, 54, 81-99.
- Wise, S. L. *Using partial orders to determine unidimensional item sets appropriate for item response theory*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, March 1982.

Author's Address

Send request for reprints or further information to Steven L. Wise, Department of Guidance and Educational Psychology, Southern Illinois University, Carbondale IL 62901, U.S.A.