DOCUMENT RESUME

ED 137 374 TM 006 178

AUTHOR Dziuban, Charles D.; Shirkey, Edwin C.
TITLE Sampling Adequacy and the Semantic Differential.
PUB DATE [Apr 77]
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Correlation; *Factor Analysis; Interaction; Matrices; *Psychometrics; *Semantic Differential
IDENTIFIERS *Measure of Sampling Adequacy

ABSTRACT

The Kaiser Measures of Sampling Adequacy (MSA) were derived for a typical six-concept Semantic Differential. The overall indices indicated that both concept and total correlation matrices would lead to comparable decisions regarding the psychometric quality of the sample data sets. The individual MSA's, however, revealed considerable variability for some scales placing several in a range which would make them suspect in a psychometric sense. It was recommended that the concepts of psychometric adequacy be used in determining the efficacy of one's Semantic Differential data for factor analytic procedures. (Author)

SAMPLING ADEQUACY AND THE SEMANTIC DIFFERENTIAL

Charles D. Dziuban

and

Edwin C. Shirkey

Florida Technological University

Paper Presented at the Annual Meeting of the
American Educational Research Association

New York City

April 4-8, 1977

Abstract

The Kaiser Measures of Sampling Adequacy (MSA) were derived for a typical six-concept Semantic Differential. The overall indices indicated that both concept and total correlation matrices would lead to comparable decisions regarding the psychometric quality of the sample data sets. The individual MSA's, however, revealed considerable variability for some scales placing several in a range which would make them suspect in a psychometric sense. It was recommended that the concepts of psychometric adequacy be used in determining the efficacy of one's Semantic Differential data for factor analytic procedures.

Of the many assessment techniques developed in the past decades, the Semantic Differential (SD) has been incorporated into an astounding number of studies. The device presents some attractive alternatives to more traditional instruments in that for one administration the responses yielded by the SD produce a large amount of data. Customarily the semantic space is portrayed as three-dimensional and Euclidean in nature. Through effective selection of scales and/or concepts, the instrument may be tailored to the specific situation encountered by the investigator and with proper construction, completed with a minimum of time and effort. These with additionally favorable considerations have made it possible to find in almost every behavioral research journal at least some studies for which the SD had been the fundamental data collection device. See for example: Aiken, 1970; Aiken, 1972; Aiken and Dreger, 1961; Anttonen, 1969; Divesta, 1966; Dutton, 1956; Husek and Wittrock, 1962; Jenkins and Suci, 1958; McCallon and Brown, 1971; Neale, 1969; Neale and Proshek, 1967; Tanaka, Oyama and Osgood, 1963; Yamamato, Thomas and Karns, 1969; Yamamato, Thomas, and Weirsma, 1969.

In addition, it has been subject to a large number of studies regarding its psychometric properties. This is best attested to by the recent volume edited by Snider and Osgood (1969). Many investigators such as Osgood and Succi (1955) have been concerned with factorial methods of structuring Semantic Differential dimensionality. Miron and Osgood (1966) noted that the literature of Semantic Differential research over a period of twelve years indicated that the scale by concept interaction was often documented but found little evidence for person by scale structure interaction. Deutschman (1959) reported that in factor analytic cases studied there were isolated

three factors--evaluation, potency, and activity. McKie and Foster (1972) proposed a model in which evaluative scales were determined through factor analytic methods with subsequent determination of concept clusters achieved through the analysis of person by concept data. Komarita and Bass (1967) reported finding three evaluative factors while Bashook and Foster (1973) argued that there was only one effective factor in those data. Moss (1960) found various sets of concepts to alter scale factor patterns, while Peabody (1967) indicated that the typically found evaluative factor resulted from a confounding of evaluative and descriptive dimensions.

Tanaka and Osgood (1965) identified a good deal of cross-cultural pattern stability in the Semantic Differential. Rosenbaum, Rosenbaum, and McGinnis (1971) indicated that patterns differed considerably with varying concepts. Recently Maguire (1973) in a comprehensive review of Semantic Differential methodology recommended that the method of principal components be abandoned and replaced with the factor analytic or image model. From these results it seems clear that factoring procedures have played a large part in structuring the dimensionality of Semantic Differential data.

Recently Kaiser (1970) announced the development of a Measure of Sampling Adequacy (MSA) which is intended to assess the degree to which a set of variables under consideration comprise an adequate sample (psychometric) from the domain of interest. The index may be used as the basis of a decision rule for determining whether a given sample correlation matrix should be factored in an exploratory sense (Cerny and Kaiser, In Press). The fundamental theory underlying the Measure (Guttman, 1953) is that as the sample data conform to factor analytic tenets, the matrix of their correlations should have

an inverse R^{-1} which approaches a diagonal (R^{-1} = Diag). Kaiser and Rice (1974) using this property defined the MSA as a function of the anti image correlation matrix $Q = SR^{-1}S$ where $S^2 = (\text{Diag } R^{-1})^{-1}$ and the observed sample correlation matrix (R):

$$MSA = \frac{\sum\sum_{j \neq k} r^2_{jk}}{\sum\sum_{j \neq k} r^2_{jk} + \sum\sum_{j \neq k} q^2_{jk}}$$

The index lies between zero and one ($0 \leq MSA \geq 1$) with values increasing with the psychometric quality of the data. The present accepted calibration for MSA is as follows:

In the .90's - marvelous
In the .80's - meritorious
In the .70's - middling
In the .60's - mediocre
In the .50's - miserable
Below .50 - unacceptable

A similar measure may be defined for individual variables.

$$MSA(j) = \frac{\sum_{k \neq j} r^2_{jk}}{\sum_{k \neq j} r^2_{jk} + \sum_{k \neq j} q^2_{jk}}$$

The individual index is intended to assess whether a particular variable represents the domain of interest and thus should be included in the sample data set.

While holding the others constant, the MSA appears to improve as (a) the number of variables increases, (b) the (effective) number of factors decreases, (d) the number of subjects increases, and (d) the general level of correlation

increases. Recent studies with the MSA (Dziuban and Shirkey, 1974A; Dziuban and Shirkey, 1974B; Shirkey and Dziuban, 1976) have shown that it signals random variables in a data set and that it would readily guard against data where the population correlation matrix was an identity (R = I). Monte carlo studies of the index (Cerny and Kaiser, In Press; Meyer, Kaiser, and Cerny, In Press; Dziuban and Shirkey, 1976) have shown it to be most influenced by the number of variables P.

The purpose of this study was to assess variations in the sampling adequacy of Semantic Differential scales when used with differing concepts. In a general sense it was a study of the psychometric interaction of scales and concepts.

Procedures

A six concept Semantic Differential was administered to a sample of five hundred fifteen public school teachers in the Central Florida area. The concepts Death, Hero, Quicksand, Success, Gentleness and White Rosebuds were among those used by Osgood, Succi and Tannenbaum (1957) to develop the instrument. Thirty scales (evaluation, potency, and activity) were selected for use from the thesaurus published by the developers. The order and polarity of scale presentation was randomly assigned. The order of concept presentation was varied randomly from a list of twelve possibilities. The scale intercorrelation matrices were derived for each of the concepts as well as for the strung out or collapsed matrix. The overall and individual Measures of Sampling Adequacy (MSA) were computed for each of the seven matrices. Although not reported in this paper, the rescaled image components were computed for

each matrix. The number of components retained (q) was equal to the number of eigenvalues of R greater than one and was a quantity of interest to this study.

Results

The overall Measures of Sampling Adequacy (MSA) and number of components are presented in Table I.

Insert Table I About Here

It may be observed that, had the MSA criterion been used exclusively as a factorability decision rule, all matrices would have been evaluated as "appropriate." The highest overall MSA (.98) was obtained for Hero and the lowest Death (MSA = .85). It is worthy of note that the total sample did not produce the largest value (MSA = .88), although the effective number of subjects appeared to be 3090. Further it may be observed that there was substantial variability in the number of retained components. The largest number (eight) was obtained for Quicksand and the lowest (three) for Hero, while the total sample produced six components.

The individual MSA's are presented in Table II.

Insert Table II About Here

If one were to view as suspect variables with values below .70, a series of psychometric interactions may be observed. A summary of those scales is presented in Table III.

Insert Table III About Here

For at least one concept eight (20.0%) of the original thirty scales produced an MSA which would be considered very suspect. Five of those were related to the originally defined activity factors with one each for potency and evaluation. None of the scales, however, within the context of the strung out matrix produced an individual MSA which normally would mandate further scrutiny. Of the six concepts, Death and White Rosebuds produced the largest number of unacceptable values, four and five respectively, while Quicksand produced three. Success and Gentleness produced one low value each while Hero was the only concept which failed to yield any clearly unacceptable MSA levels.

Discussion

Through the use of some kind of factoring procedure, it has been and still is common practice to assess the underlying dimensionality of Semantic Differential data. Much of this work has been exploratory in nature, although often the intent has been to retrieve the evaluation, potency, and activity dimensions. Because of its three-dimensional nature, exploratory factor analysis of Semantic Differential data has presented complex analysis problems. If one wishes to analyze a two-dimensional array of scales, however, two primary options are available. The first, which is rarely done, is to analyze the scales for each concept separately and subsequently compare the similarity of the factors. The second method involves deriving the factors for the strung out matrix by collapsing the concepts and forcing non-

independent observations. We have assessed the overall psychometric adequacy of the one data set in both situations and found it to be acceptable. In addition, individual indices for the strung out matrix showed all of the variables to be at least minimally adequate. From this one might assume that they could proceed with the proposed factoring procedures.

The results of the concept factoring, however, indicated that certain precautions should be taken in these circumstances. The change in psychometric adequacy of some scales, when used with different concepts, may indicate a change in domain relatedness. This becomes evident when one "dissects" the Semantic Differential. When the individual MSA's for the overall matrix were computed the scale pair Weak-Strong produced a value (.79) which would by present calibration standards lead one to decide that it had an acceptable domain relatedness. It did, however, produce decidedly inferior values for two of the concepts (.46, .59). The effect of collapsing the concepts was to obscure the inferior MSA's. Since this occurred with eight scales, it seems worthy of consideration.

The results of this study suggest that, depending on the concept, application of a common decision rule ($\lambda \geq 1$) to Semantic Differential data will produce largely disparate numbers of components. With as many as eight and as few as three dimensions found in the separate concepts the total of six in the overall matrix suggested "average" components. Apparently through the same decision rule we have over or under factored each one of the concepts. The data from the Semantic Differential presents a natural three-dimensional data box. Analyzing the components separately for each concept tends to ignore the basic structure of the instrument as does collapsing the data over

concepts. An ideal analysis should involve some form of a three-mode procedure such as the one proposed by Tucker (1972).

Should one, however, choose a conventional analysis strategy, the overall and individual Measures of Sampling Adequacy will be useful in reaching appropriate decisions regarding the quality of one's data set. The overall measures provide important information as to whether the data should be factored at all. The individual measures might provide information as to which variables should be included in subsequent data collection and analysis strategies. A variable which is universally poor is probably a good candidate for deletion. On the other hand, scales with erratic sampling adequacy characteristics may be well suited for some concepts and not for others. This scale by concept interaction has long been recognized so we re-emphasize that scale selection is a critical step in the assembly of a particular Semantic Differential. The individual MSA's might be further helpful in this context since with the SD one is factoring single scales with their presumed low reliability instead of batteries of tests. Accordingly, we feel the concept of psychometric sampling adequacy should provide a helpful guide to the analysis of Semantic Differential and recommend that one scrutinize his/her data by examining the associated overall and individual MSA's.

References

Aiken, L. Attitudes Toward Mathematics. Review of Educational Research, 1970, 40, 551-596.

Aiken, L. Research on Attitudes Toward Mathematics. Arithmetic Teacher, March, 1972, 229-234.

Aiken, L. & Dreger, R. The Affect of Attitude on Performance in Mathematics. Journal of Educational Psychology, 1961, 52, 19-24.

Anttonen, R. A Longitudinal Study in Mathematics Attitude. Journal of Educational Research, 1969, 62, 467-471.

Bashook, P. G., & Foster, S. F. How Many E's Are There? - A Critical Analysis of Problems Concerning Determination of Evaluative Factors of Semantic Differential Scales. A paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, February 25-March 1, 1973.

Cerny, B. A., & Kaiser, H. A. A Study of a Measure of Sampling Adequacy for Factor-Analytic Correlation Matrices. Multivariate Behavioral Research, (In Press).

Deutschman, P. The Semantic Differential, Its Uses and Abuses. Public Opinion Quarterly, 1959.

Divesta, F. The Test Retest Reliability of Children's Ratings on the Semantic Differential. Educational and Psychological Measurement, 1966, 109, 205-229.

Dutton, W. Attitudes of Junior High School Pupils Toward Arithmetic. School Research, January, 1956, 18-22.

Dziuban, C. D., & Shirkey, E. C. On the Psychometric Assessment of Correlation Matrices. American Educational Research Journal, 1974A, 11, 211-216.

Dziuban, C. D., & Shirkey, E.C. When Is a Correlation Matrix Appropriate for Factor Analysis? Some Decision Rules. Psychological Bulletin, 1974B, 6, 358-361.

Dziuban, C. D., & Shirkey, E. C. An Investigation of Some Distributional Characteristics of the Measure of Sampling Adequacy. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.

Guttman, L. Image Theory for the Structure of Quantitative Variates. Psychometrika, 1953, 18, 277-296.

Husek, T., & Wittrock, M. The Dimensions of Attitudes Toward Teachers as Measured by the Semantic Differential. Journal of Educational Psychology, 1962, 53, 209-213.

Jenkins, S., & Suci, G. An Atlas of Semantic Profiles for 360 Words. American Journal of Psychology, 1958, 71, 688-699.

Kaiser, H. F. A Second Generation Little Jiffy. Psychometrika, 1970, 35, 401-416.

Kaiser, H. F., & Rice, J. Little Jiffy, Mark IV. Educational and Psychological Measurement, 1974, 34, 114-117.

Komorita, S. S., & Bass, A. R. Attitude Differentiation and Evaluative Scales of the Semantic Differential. Journal of Personality and Social Psychology, 1967, 6, 241-244.

Maguire, T. Semantic Differential Methodology for the Structuring of Attitudes American Educational Research Journal, 1973, 10, 295-306.

McCallon, E., & Brown, J. A Semantic Differential Instrument for Measuring Attitudes Toward Mathematics. The Journal of Experimental Education, Summer, 1971, 69-72.

McKie, D., & Foster, S. F. A General Model for Multidimensional Analysis of Semantic Differential Attitude Data. Proceedings of the Annual American Psychological Association Convention, Honolulu, 1972, 45-46.

Meyer, E. P., Kaiser, H. F., & Cerny, B. A. MSA for a Special Spearman Matrix. Psychometrika, (In Press).

Miron, M. S., & Osgood, C. E. The Multivariate Structure of Qualifications. In R. B. Cattell (Ed.), Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally, 1966.

Moss, C. S. Current and Projected Status of Semantic Differential Research. Psychological Record, 1960, 10, 47-54.

Neale, D. The Role of Attitudes in Learning Mathematics. Arithmetic Teacher, December, 1969, 16.

Neale, D., & Proshek, J. School Related Attitudes of Culturally Disadvantaged Elementary School Children. Journal of Educational Psychology, 1967, 58, 238-244.

Osgood, C. E., & Succi, G. J. Factor Analysis of Meaning. Journal of Experimental Psychology, 1955, 50, 325-328.

Osgood, C. E., Succi, G. J., & Tannenbaum, P. H. The Measurement of Meaning. Urbana: University of Illinois Press, 1957.

Peabody, D. Trait Inferences: Evaluative and Descriptive Aspects. Journal of Personality and Social Psychology, 1967, 7 (Monogr. Suppl. 4, Whole No. 644).

Rosenbaum, L. L., Rosenbaum, W. R., & McGinnies, E. Semantic Differential Factor Structure Stability Across Subject, Concept, and Time Differences. Multivariate Behavioral Research, 1971, 6, 451-469.

Shirkey, E. C., & Dziuban, C. D. A Note on Some Sampling Characteristics of the Measure of Sampling Adequacy (MSA). Multivariate Behavioral Research, 1976, 11, 125-128.

Snider, J. R. & Osgood, C. E. (Eds.). The Semantic Differential Technique: A Book of Readings. Chicago: Aldine Publishing Company, 1969.

Tanaka, Y., & Osgood, C. E. Cross-Culture, Cross Concept, and Cross Subject Generality of Affective Meaning Systems. Journal of Personality and Social Psychology, 1965, 2(2), 143-153.

Tanaka, Y., Oyama, T., & Osgood, C. A Cross-Culture and Cross Concept Study of the Generality of Semantic Spaces. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 392-405.

Tucker, C. R. Some Mathematical Notes on Three-Mode Factor Analysis. Psychometrika, 1972, 37, 2-28.

Yamamoto, K., Thomas, E., & Karns, E. School Related Attitudes in Middle-School Age Students. American Educational Research Journal, March, 1969, 191-206.

Yamamoto, K., Thomas, E., and Weirsma, S. Dimensions of School Related Attitudes in Middle-School Age Students. Psychology in the Schools, 1969, 10, 375-378.

Table I

Overall MSA's and Number of Components

Concept	MSA	# of Components
Quicksand	.88	8
Death	.85	7
White Rosebuds	.92	7
Hero	.95	3
Gentleness	.92	4
Success	.93	5
Total	.88	6

Table II

Individual Measures of Sampling Adequacy*

	White Rosebuds	Hero	Gentleness	Success	Quicksand	Death	Total Matrix
Passive/Active	59	94	67	90	86	68	78
Soft/Hard	91	93	94	91	86	87	75
Complete/Incomplete	93	95	96	94	79	86	92
Excitable/Calm	79	96	85	94	92	68	89
Good/Bad	93	96	91	95	91	90	89
Weak/Strong	46	92	88	92	85	59	79
Awkward/Graceful	92	96	93	92	84	83	88
Fast/Slow	89	97	91	64	50	65	80
Tenacious/Yielding	91	94	89	91	90	71	90
Free/Constrained	96	96	96	72	92	87	93
Pleasurable/Painful	92	95	94	95	90	91	91
Energetic/Inert	67	97	88	96	68	87	89
Simple/Complex	91	93	87	93	71	78	80
Serious/Light	91	96	80	95	94	87	88
Cold/Hot	46	93	89	83	87	83	77
Thick/Thin	89	92	81	92	93	75	91
Beautiful/Ugly	94	96	94	94	92	86	91
Severe/Lenient	92	81	87	90	93	93	91
Easy/Difficult	95	90	94	91	90	86	90
Large/Small	90	96	88	95	92	81	91
Emotional/Unemotional	90	95	95	94	88	85	84
Untimely/Timely	91	94	92	94	81	87	90
Light/Heavy	93	90	94	93	89	89	80
Cruel/Kind	93	95	90	85	83	89	84
Dead/Alive	92	95	94	93	75	90	87
Moving/Still	63	97	89	95	89	88	90
Meaningful/Meaningless	94	97	92	95	62	75	91
Masculine/Feminine	96	96	84	86	86	78	89
Unimportant/Important	95	93	92	92	73	72	81
Positive/Negative	96	95	94	95	89	83	94

*Decimals Omitted

MSA < .70 Underlined

Table III

Summary of the Scales
for Which Deficient MSA's Were Found

	White Rosebuds	Hero	Gentleness	Success	Quicksand	Death
Passive/Active	59	--	67	--	--	68
Excitable/Calm	--	--	--	--	--	68
Weak/Strong	46	--	--	--	--	59
Fast/Slow	--	--	--	64	50	65
Energetic/Inert	67	--	--	--	68	--
Cold/Hot	46	--	--	--	--	--
Moving/Still	63	--	--	--	--	--
Meaningful/Meaningless	--	--	--	--	62	--