Classification of Community Hospitals by Scope of Service: Four Indexes

by Mary Edwards, Jon D. Miller, and Rex Schumacher

Four indexes are presented for classifying short-term nonfederal general hospitals by the scope of service they provide. The indexes, constructed by the application of Guttman scaling to data from 5439 hospitals, are tested for cohesiveness and unidimensionality and their relation to hospital expenses and staffing is examined. The usefulness of the indexes for classifying hospitals and as stratification variables is discussed.

The nature of hospital care differs significantly from institution to institution. These differences are usually related to the mixture of specific services comprising a unit of hospital care or, more generally, to the scope of services available to patients. While differences in the scope of service have been noted by a number of investigators, few attempts have been made to systematically control for these differences.

Historically, the major variables used to classify hospitals have been type of service, type of control, average length of stay, and number of inpatient beds. A combination of the first three of these variables defines the general category of community hospitals: short-term, nonfederal hospitals offering general and selected special services. With federal units, psychiatric and tuberculosis hospitals, and all other long-term institutions excluded, the category of community hospitals is more homogeneous than the total universe of hospitals.

Within the category of community hospitals, the major—and almost the only—variable employed for analytic purposes has been hospital size, measured by the number of inpatient beds. The almost total reliance on size, especially in cost studies, has led to a number of confusing and contradictory findings pertaining to possible economies of scale. Lave [1] and Hefty [2] have provided useful summaries of these studies.

In one of the few studies attempting to account for differences in scope of services, Carr and P. Feldstein [3] grouped hospitals according to the number of facilities and services and performed multiple regression analyses on each group. This procedure provided a rough control for differences in scope of service, but because of the relatively wide range in number of services within categories, considerable differences remained.

Address communications and requests for reprints to Mary Edwards, Ph.D., Division of Social Sciences, SUNY–College at Purchase, Purchase, N.Y. 10577.

In another study of economies of scale, Berry [4] identified 40 groups of hospitals with identical services and facilities and employed simple bivariate linear regressions for his analysis. The grouping of hospitals with identical facilities and services effectively controlled for differences in the scope of services, but the small number of cases in each category vitiated much of the subsequent analysis, certainly ruling out more sophisticated multivariate models. The identical-service approach was also used by Francisco [5], but matching eliminated two-thirds of community hospitals from analysis.

In another analysis, Francisco grouped hospitals with the same number of services (but not necessarily identical services) and repeated the regression analysis. While this approach offers substantial utility and simplicity, again the grouping reduced the number of hospitals included in the analysis.

Other investigators—principally M. Feldstein [6] and Cohen [7]—have discarded the scope of services approach and have attempted to develop new measures of differences in composition of hospital care. Working with British hospitals, M. Feldstein developed nine diagnostic-therapeutic categories for patients and used the proportion of cases in each category to classify hospitals. Cohen introduced a technique in which the number of various hospital services delivered is weighted by their relative average cost. While both approaches appear to offer the potential for sophisticated analysis, they require infrequently available data that can be collected only with substantial expenditures of time and money.

In addition, Cohen's technique is useful primarily for cost analysis and does not provide a general classification of hospitals. By contrast, the indexes described in this article should serve a wide variety of purposes, including classification as well as cost analysis. The simpler measures of service level proposed here may be adequate at this point for a variety of research purposes.

Basis of the Indexes

Hospitals appear to fit a model of cumulative growth with regard to scope of services. Clearly, a hospital does not provide a sophisticated service such as an organ bank without first providing simpler services such as a pharmacy. This growth model is supported by Kaluzny et al. [8], who applied scaling techniques to five outpatient services in a study of addition of services. The ideal index of scope of services should reflect the cumulative nature of hospital services, with a higher score on the index indicating that a hospital offers not only more services but more sophisticated services than a lower-scoring hospital. For this reason, the general service index and the three subindexes developed here—acute inpatient care, long-term care, and outpatient care—were based on a model of cumulative scaling, in which services are assumed to be acquired in order of increasing sophistication and in which the latest acquisitions represent a greater scope of service than earlier acquisitions.

Guttman Scalogram Technique

The specific model is the Guttman scale [9], a technique used originally in attitude research. The most important characteristics of the Guttman scale are unidimensionality and the cumulative quality of the items. Unidimensionality simply means that all the items are measuring the same variable. The cumulative nature of the scale means that different items measure different degrees of that variable. The items can be arranged in order by the degree of the variable they possess, with items high on the scale subsuming items lower on the scale. For example, it can be assumed that an organ bank represents a higher level of service than a pharmacy and that a hospital with an organ bank will have a pharmacy. Because the items are scaled in this fashion, the hospital's total score on the Guttman scale is a predictor of the services available in that hospital.

The aim of scalogram analysis is to determine whether a set of respondents (hospitals) and a set of items (services) can be logically ordered together to form a unidimensional scale. With real data, an errorless Guttman scale will seldom be found. The problem is to decide whether the data are close enough to the model to justify ignoring the errors that occur. Unfortunately, there is no single measure that indicates whether a set of items conforms to the Guttman model. Guttman's original criterion, the coefficient of reproducibility [9], is essentially the proportion of correct responses found. If a scale is valid, a hospital's total score should predict exactly which services the hospital provides. The services actually available in the hospital are compared with the predicted services, and each disagreement contributes to the error score. Guttman somewhat arbitrarily decided that 10 percent error is allowable in an acceptable scale.

A number of problems are associated with this coefficient. Extreme respondents or items (those with all "yes" or all "no" responses) inflate the coefficient, since they cannot contribute error. In addition, the reproducibility for an item is seldom less than the proportion of respondents giving the modal response [10], and the chance reproducibilities are quite high for scales with a small number of items [11,12]. These difficulties can be met in part by using a relatively large number of items in scaling and by choosing items that approximate a 50 percent distribution among the respondents. The null hypothesis that an observed distribution arose by chance can be tested with some accuracy [13–15], but this measures only the items' departure from total independence, and the Guttman hypothesis of perfect scalability remains untested. In fact, high reproducibilities have been obtained for data in which two totally unrelated subsets of items were combined [16]. Obviously, reproducibility is not equivalent to unidimensionality.

Cohesiveness and unidimensionality can be better tested by examining the relation of each item to the others and to the scale as a whole. If items are cohesive, they will be positively associated with the total scale score [17]. A positive correlation between an item and the total score would mean that higher-scoring hospitals are more likely to have the service than lower-scoring hospitals.

The unidimensionality of the items can be examined through correlations among individual items. A positive correlation between items (a hospital with service *A* is likely to have service *B*) would indicate that the items are measuring the same variable. While correlations among items should be positive, the literature does not suggest a minimum acceptable value for assuring unidimensionality. Further testing of this quality can be carried out through a modified factor analysis based on Lawley's maximum likelihood method [16]. This analysis indicates the extent to which all the items are related to one another and therefore to a single factor.

The approach taken in developing the four indexes presented in this article was first to test proposed scale items for cohesiveness and unidimensionality and then to test for fit to the Guttman model by comparison of chance and actual reproducibilities.

Index Construction

Testing Cohesiveness and Unidimensionality

Data on the services of 5439 community hospitals were taken from the 1969 annual survey of registered hospitals conducted by the American Hospital Association [18]. The hospital auxiliary was eliminated because it is not a patient service, and the outpatient psychiatric and organized outpatient departments were eliminated because of problems of definition and measurement in the annual survey. Hospitals with either a full-time or a part-time pharmacist were categorized as having a pharmacy, so the original pharmacy category was collapsed. These data then indicated the presence or absence of each of 31 services. The total number of services provided in each hospital was obtained, and the item-total correlations (phi coefficients) were calculated for each service. These correlations were arranged in descending order, and the number of services was successively reduced by eliminating the least cohesive items and recalculating correlations among the remainder.

As services were eliminated a 17-item scale appeared to be optimal, since further eliminations did not improve the correlations. Two services that had strong theoretical appeal, however—inpatient renal dialysis and organ bank—were eliminated by this procedure. Because their relatively low item-total correlations were related more to their low frequency than to a lack of cohesiveness, they were reincluded in the general index. Neither substantially altered scalability or the unidimensionality of the scale. Table 1 shows the final general index of services arranged in their Guttman scale order, with their final item-total correlations.

Of the three subindexes created, the acute-care index (which is simply the acute-care component of the general index) is most directly related to the general index. When the three nonacute services included in the general index—social work, emergency psychiatry, and inpatient psychiatry—were dropped and item-total correlations recalculated, coefficient values did not change substantially.

Table 1. Guttman Scale Order and Item-Total Correlations of Services in Final General Index

Service	Guttman scale order	Item-total correlation (phi coefficient)
None	1	
Pharmacy	2	.41
Postoperative recovery	3	.49
Physical therapy	4	.58
Inhalation therapy	5	.58
Intensive care	6	.67
Histopathology	7	.68
Intensive cardiac care	8	.54
X-ray therapy	9	.66
Radioisotope laboratory	10	.74
Electroencephalography	11	.67
Radium therapy	12	.68
Social work*	13	.58
Emergency psychiatry*	14	.47
Inpatient psychiatry*	15	.52
Occupational therapy	16	.48
Inpatient renal dialysis	17	.36
Cobalt therapy	18	.54
Open-heart surgery	19	.46
Organ bank	20	.22

* Acute-care index consists of 17 items remaining when three services marked with asterisk are deleted; item-total correlations for acute-care index do not differ appreciably from those shown.

The number of items available for inclusion in the outpatient and long-term-care indexes was quite limited—the original list of services included only seven long-term-care items and nine outpatient items. As mentioned earlier, scales with small numbers of items tend to be unreliable, with high chance reproducibilities and spurious scaling. Because of this problem, the process of item elimination was not carried out and all the relevant items were retained in both indexes. As might have been expected, the correlations (Table 2) are rather low in comparison with the general index. These two indexes therefore are more limited in utility than the general and acute-care indexes and should be used with that understanding.

With the scale items selected and evaluated for cohesiveness, the indexes were tested for unidimensionality. As explained earlier, the correlations among items constitute a rough measure of unidimensionality. For all four indexes these correlations were all positive and therefore fit the expected pattern.

A more precise testing of the unidimensionality of the four indexes was carried out with the modified factor analysis mentioned previously [16]. The strength of the association between individual scale items and the proposed factor—scope of service—is indicated by the factor loadings. For both the general and the acute-care indexes the services included had relatively high loadings on

Table 2. Guttman Scale Order and Item-Total Correlations of Services in Final Outpatient and Long-term-care Indexes

Service	Guttman scale order	Item-total (phi coefficient) correlation
OUTPATIENT INDEX		
None	1	
Emergency department	2	.11
Organized outpatient department	3	.43
Social work	4	.58
Outpatient psychiatry	5	.53
Outpatient rehabilitation	6	.43
Outpatient renal dialysis	7	.30
Home care	8	.32
Partial psychiatric hospitalization	9	.36
Family planning	10	.47
LONG-TERM-CARE INDEX		
None	1	
Physical therapy	2	.34
Inpatient psychiatry	3	.42
Inpatient rehabilitation	4	.40
Extended care	5	.21
Home care	6	.25
Partial psychiatric hospitalization	7	.35
Self-care	8	.25

the scope of service dimension. Excluding organ bank and inpatient renal dialysis (which had the predictably low factor loadings of .23 and .38, respectively), the factor loadings ranged from .42 to .78 for the general index items and had similar values for items in the acute-care index. As would be expected, the factor loadings for items in the other two subindexes were substantially smaller. They ranged from .12 to .70 for the outpatient index and from .24 to .62 for the long-term care index, indicating only moderate association between the individual items and the scope of service dimension.

While the factor loadings indicate the contribution of the individual items to the factor, the measurement of single-factoredness hinges on the relation between the variance explained and the residual variance. The Wolins index describes this relation: the index has a value of 1.00 when the first factor leaves only the residuals expected by chance in a one-factor case and a value of zero if the first factor explains no variance. A value of .75 has been recommended as the lower bound of acceptable unidimensionality [16]. For all four indexes the Wolins value was greater than this minimum, indicating that one factor accounted for 87 percent or more of the variance in the original correlation matrix for each scale.

In sum, all items included in the four indexes were tested for cohesiveness and unidimensionality. Items in the general and acute-care indexes proved to be quite cohesive with their respective scales. Although this was less true of

Table 3. Coefficients of Reproducibility for Real and Random Data for All Four Indexes

Service index	Coefficient of reproducibility ($N = 5439$)	
	Random data	Real data
General	.72	.87
Acute care	.73	.87
Outpatient	.83	.92
Long-term care ...	.85	.92

items in the outpatient and long-term-care indexes, the item-total correlations in all cases were positive, indicating at least minimum cohesiveness for all four indexes. In terms of unidimensionality all four indexes were found to be above the acceptable minimum level. The scale items appear to be measuring a single variable, here called "scope of service." These results provide empirical support for the original assumption of unidimensionality as a basis for the indexes of scope of service.

Testing Scalability

With cohesiveness and unidimensionality assured, the four indexes finally were tested for fit to the Guttman model, using Guttman's original coefficient of reproducibility [9]. The coefficients were calculated for the real data and for random data for all indexes and the two sets of results compared, as shown in Table 3. For all four indexes the amount of error in the real data was relatively low, ranging from 13 percent for the general and acute-care indexes to 8 percent for the outpatient and long-term-care scales. In contrast, the random data contained approximately twice as many errors as the real data. Clearly, the real data more closely approximate the Guttman model, producing adequate scales with tolerable amounts of error.

In terms of both the internal consistency of the items and the scalability of the item sets, the general and acute-care indexes appear to be completely acceptable scales. The outpatient and long-term-care indexes, while reaching acceptable levels of unidimensionality and scalability, should be used with more caution because of the smaller number of scale items and the problems of measurement in several of the items.

Characteristics of Indexes

Hospital Subgroups

Obviously, measures of scope of service that are valid for various subcategories of community hospitals would be considerably more useful than measures that are relevant only for one or another type. The indexes have been developed on the basis of data for all community hospitals regardless of type of control.

Because control provides another important classification of hospitals, the tests for cohesiveness, unidimensionality, and reproducibility were run for community hospitals grouped in the three major control subcategories: nongovernmental nonprofit, or voluntary, hospitals; nongovernmental for-profit, or proprietary, hospitals; and governmental nonfederal, or hospitals run by state and local governments. The three separate control categories differed little from one another in terms of cohesiveness, unidimensionality, or scalability of the four indexes.

Distribution of Hospitals on Index Scales

The distribution of hospitals by their Guttman scores on each of the four indexes, shown in Table 4, illustrates the levels of service provided by commu-

Table 4. Distribution of Community Hospitals by Index Score for Each Index

Score*	Hospitals ($N = 5439$)		Score*	Hospitals ($N = 5439$)	
	Number	Percent		Number	Percent
		GENERAL INDEX			
1	382	7.0	11	227	4.2
2	582	10.7	12	203	3.7
3	578	10.6	13	209	3.8
4	534	9.8	14	198	3.6
5	450	8.3	15	144	2.7
6	401	7.4	16	142	2.6
7	319	5.9	17	107	2.0
8	293	5.4	18	76	1.4
9	280	5.2	19	50	0.9
10	235	4.3	20	29	0.5
		ACUTE-CARE INDEX			
1	417	7.7	10	254	4.7
2	610	11.2	11	290	5.3
3	610	11.2	12	250	4.6
4	551	10.1	13	239	4.4
5	457	8.4	14	169	3.1
6	415	7.6	15	117	2.2
7	323	5.9	16	72	1.3
8	323	5.9	17	37	0.7
9	305	5.6			
	OUTPATIENT INDEX			LONG-TERM-CARE INDEX	
1	495	9.1	1	1959	36.0
2	2445	44.9	2	1935	35.6
3	1173	21.6	3	840	15.4
4	525	9.7	4	411	7.6
5	310	5.7	5	196	3.6
6	211	3.9	6	73	1.3
7	140	2.6	7	18	0.3
8	79	1.5	8	7	0.1
9	52	0.9			
10	9	0.2			

* Score of 1 means a hospital has none of services included in an index, score of 2 indicates one service, etc.

nity hospitals across the nation in 1969. (Note that an index score of 1 indicates a hospital with none of the indexed services, 10 shows a hospital with nine services, and so on.) Hospitals tend to cluster at the lower end of the scale on all four indexes. One-third of community hospitals had scores of 4 or less on the general index, while less than 1 percent of hospitals had the maximum index score of 20. A similar distribution occurs for the acute-care index. The distribution of hospitals on the outpatient index was heavily weighted toward the lower end of the scale, with slightly more than one-half of reporting community hospitals falling in the first two levels—that is, offering either no outpatient service or only an emergency department. Only 1 percent of all hospitals offered either eight or nine of the outpatient services included in the index. The distribution of hospitals on the long-term-care index was even more bottom-heavy.

Since the indexes are designed for use across time, the clustering of hospitals toward the low end provides growth potential for the measures. As hospitals acquire more services their index scores will increase accordingly. Since so few hospitals now score at the highest levels of the indexes, growth in scope of services can be charted for some time to come without the risk of imposing an artificially low ceiling on index values.

Relation to Hospital Expenses and Staffing

The basic reason for developing the indexes was to find a simple variable that would reflect scope of service for use in studying differences in expense and staffing patterns. The relationship between scope of service and total hospital expenditures was examined by use of a multiple regression similar to that employed by Carr and Feldstein [3]. While a full report of these findings must await a separate treatment, the preliminary results indicate that it is possible to explain 90 percent of the variance by use of only the general index, the number of adjusted patient days, and the second-degree polynomials of these two variables (see Table 5). In fact, once scope of service and volume of service units have been taken into consideration, the influence of urban or rural location, the existence of a nursing school, and the operation of residency programs have little effect on total expenditure. The influence of these factors may be indirect in that they encourage a broader scope of services, but these preliminary data sug-

Table 5. Summary of Stepwise Multiple Regression with Total Expenditures as Dependent Variable

Independent variable	Variance explained (R^2)	Increase in R^2	F ($p < .05$)
General index	.563		5367
(General index)^2	.673	.111	1412
Adjusted patient days	.876	.202	5525
(Adjusted patient days)^2	.890	.015	574
Population density (SMSA)	.899	.009	369
Residency program	.901	.002	74
Nursing school	.902	.001	27

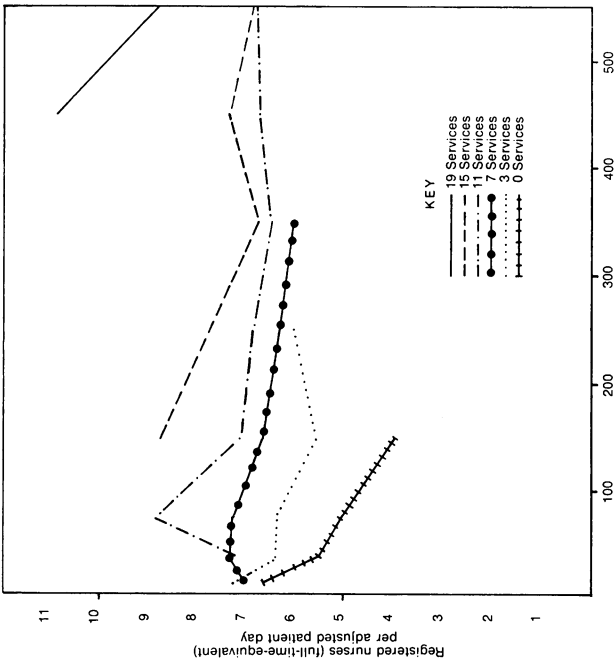

Fig. 2. Average number of full-time-equivalent registered nurses, by hospital size and selected general index levels, 1969.

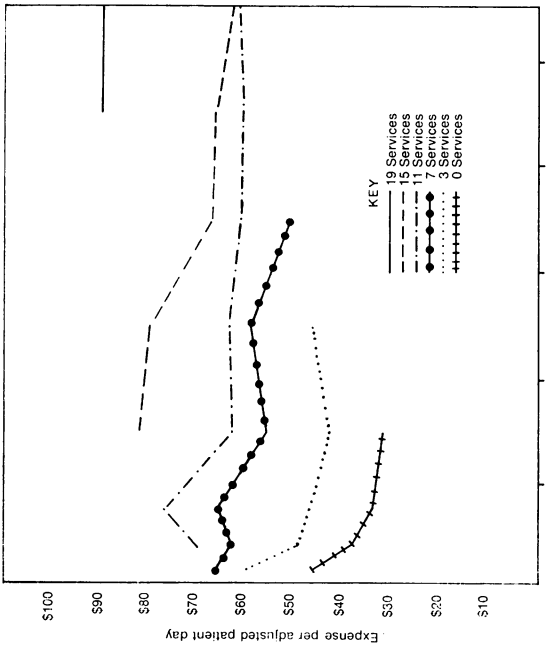

Fig. 1. Average expense per adjusted patient day, by hospital size and selected general index levels, 1969.

gest that it is the broader scope of services and not an urban location or a teaching program that is associated with higher expenditures.

As shown in Fig. 1, a similar result can be observed by examining the average expense per adjusted patient day for various levels of the indexes. Working with the general index again, the average expense per adjusted patient day is higher for hospitals at each higher level of the index, but there is a slight tendency for it to decrease with increasing hospital size, suggesting a small return to scale within each index level.

A similar association was found between the number of full-time-equivalent registered nurses and the general service index: over 80 percent of the variance was explained with the same small set of variables. As shown in graphic form in Fig. 2, staffing level also increases with index score but tends to decline with increasing hospital size, again suggesting a small economy of scale at a given service level. These patterns were also generally found to hold within different hospital control groups.

From these preliminary results, it appears that these indexes will prove to be useful and parsimonious predictors of hospital expense and staffing patterns.

Discussion

Given that the internal validity of the indexes is reasonably well established, what is the major utility of these measures? First, they provide good summary variables for scope of service. The indexes allow measurement—however rough—of differences in the complexity of hospital care. In effect, one can begin to approximate product differentiation in the hospital industry. A second advantage—and perhaps the most important—is the simplicity and low cost of the indexes. Index scores can be established easily for single hospitals or selected groups, and these scores can be replicated reliably from study to study. The low cost of using the indexes should prove a particular advantage in large-scale research projects.

Third, the general index provides a good classification for studies of hospital costs and staffing and a stratification variable for drawing samples. The general index can be used in studies in which hospital costs are undifferentiated by the type of services provided. When more refined classifications are sought or multivariate classification based on type of service is called for, one or more of the subindexes might be used. In terms of stratification for sampling, preliminary work indicates that use of the general index brings about a major reduction in variance in a number of relevant variables.

A fourth advantage of the indexes developed here is their growth potential. As has been indicated, there is a tendency for hospitals to cluster toward the bottom of the scales, a tendency that is particularly pronounced for the outpatient and long-term-care indexes. Consequently they will be useful in measuring changes in service level in longitudinal studies for some years to come.

This raises the problem of the use of the indexes for historical research. The indexes are based on services reported in the 1969 annual survey of hospitals.

Earlier annual surveys do not include exactly the same items, partly because of changing definitions and partly because of the development of new hospital services. As a result, exact equivalents of the present indexes are not available for the years prior to 1969. However, the indexes have been adapted successfully for use with data from 1965 onward,[1] although it is not clear how much further back the indexes can be applied without loss of meaning. Lack of continuity is also a potential problem in the future. The continued use of the index will require the inclusion in future annual surveys of all the items included in the 1969 survey, with the same definitions.

Further investigation and validation are desirable, but the preliminary results appear to promise a new and effective variable for analyzing differences in the scope of services provided by community hospitals. It is hoped that other investigators of health services will attempt to utilize these indexes in their studies, since it is only through extensive use that the value of these measures can be established.

References

1. Lave, J. R. A review of the methods used to study hospital costs. *Inquiry* 3:57 May 1966.
2. Hefty, T. R. Returns to scale in hospitals: A critical review of recent research. *Health Serv. Res.* 4:267 Winter 1969.
3. Carr, W. J. and P. J. Feldstein. The relationship of cost to hospital size. *Inquiry* 4:45 June 1967.
4. Berry, R. E. Returns to scale in the production of hospital services. *Health Serv. Res.* 2:123 Summer 1967.
5. Francisco, E. W. Analysis of cost variations among short-term general hospitals. In H. E. Klarman (ed.), *Empirical Studies in Health Economics.* Baltimore: Johns Hopkins Press, 1970.
6. Feldstein, M. S. *Economic Analysis for Health Service Efficiency.* Chicago: Markham Publishing Company, 1968.
7. Cohen, H. A. Variations in cost among hospitals of different sizes. *Southern Econ. J.* 33:355 January 1967.
8. Kaluzny, A. D., J. E. Veney, J. T. Gentry, and J. B. Sprague. Scalability of health services: An empirical test. *Health Serv. Res.* 6:214 Fall 1971.
9. Guttman, L. The basis for scalogram analysis. In S. A. Stauffer et al. (eds.), *Measurement and Prediction.* Princeton: Princeton University Press, 1950.
10. Guttman, L. The Cornell technique for scale and intensity analysis. *Educ. & Psychol. Measurement* 7:247 Summer 1947.
11. Guttman, L. Problems of reliability. In S. A. Stauffer et al. (eds.), *Measurement and Prediction.* Princeton: Princeton University Press, 1950.
12. Green, B. F. Jr. Attitude measurement. In G. Lindzey (ed.), *Handbook of Social Psychology.* Reading, Mass.: Addison-Wesley, 1954.
13. Sagi, P. C. A statistical test for the significance of a coefficient of reproducibility. *Psychometrika* 24:19 March 1959.
14. Schuessler, K. E. A note on statistical significance of scalogram. *Sociometry* 24:312 September 1961.
15. Chilton, R. J. Computer generated data and the statistical significance of scalogram. *Sociometry* 29:175 June 1966.

[1] Several methods for establishing index scores for earlier years have been developed by David Peterson and Bruce Neuman of the Northwestern University School of Management.

16. Morrison, D. G., D. T. Campbell, and L. Wolins. A Fortran program for evaluating internal consistency and single-factoredness in sets of multilevel attitude items. Vogelback Computing Center, Northwestern University, 1967.
17. Guilford, J. P. *Psychometric Methods*, rev. ed. New York: McGraw-Hill, 1954.
18. *Hospitals, J.A.H.A.* Aug. 1, 1970, Part Two, Guide Issue.