measurement." Journal of Personality and Social Psychology 1:28–42.
Rosenthal, Robert
1966 Experimenter Effects in Behavioral Research. New York: Appleton-Century-Crofts.
Srole, Leo, Thomas S. Langner and Stanley T. Michael, *et al.*
1962 Mental Health in the Metropolis. New York: McGraw-Hill.
Wanklin, J. M., D. F. Fleming, C. W. Buck and G. E. Hobbs
1955 "Factors influencing the rate of first admission to mental hospital." Journal of Nervous and Mental Disease 121:103–116.
Wells, William D.
1963 "How chronic overclaimers distort survey findings." Journal of Advertising Research 3:8–18.

THE ASSIGNMENT OF NUMBERS TO RANK ORDER CATEGORIES *

Sanford Labovitz

University of Southern California

By both random and nonrandom assignments of numbers to rank orders (which are consistent with the monotonic nature of the categories), it is shown that ordinal variables can be treated as if they conform to interval scales. The scoring systems, of which 18 were randomly generated by a computer, resulted in negligible error when comparing any assigned scoring system with any selected "true" scoring system. Errors are determined by the Pearsonian correlation coefficient (r) *and* r^2. *The advantages of treating ordinal variables as interval are demonstrated with regard to the relation between occupational prestige and suicide. These advantages include: (1) the use of more powerful, sensitive, better developed and interpretable statistics with known sampling error, (2) the retention of more knowledge about the characteristics of the data, and (3) greater versatility in statistical manipulation (e.g., partial and multiple correlation and regression, analysis of variance and covariance, and most pictorial presentations). The computer approach to this problem does not exhaust all possibilities for assigning numbers, which partially limits the generality of the findings.*

Empirical evidence supports the treatment of ordinal variables *as if* they conform to interval scales (Labovitz, 1967).[1] Although some small error may accompany the treatment of ordinal variables as interval,[2] this is offset by the use of more powerful, more sensitive, better developed, and more clearly interpretable statistics with known sampling error. For example, well-defined measures of dispersion (variance) require interval or ratio based measures. Furthermore, many more manipulations (which may be necessary to the problem in question) are possible with interval measurement, e.g., partial correlation, multivariate correlation and regression, analysis of variance and covariance, and most pictorial presentations. The arguments presented below are general enough to apply to any ordinal scale, and perhaps with even greater confidence they apply to variables that fall between ordinal and interval, e.g., I.Q. scores and formal education (Somers, 1962:800).

* I am grateful to Robert Hagedorn, Harvey Marshall, Ross Purdy, and the referees of *ASR* for their helpful comments and critical reading of an earlier draft.

[1] Labovitz demonstrates the utility of treating ordinal variables as interval for a hypothetical problem relating two types of therapy to four subjective responses: it made me worse (−); it had no effect (0); it helped a little (+); and it helped quite a bit (++). The four ordinal responses are assigned scores ranging from highly skewed (e.g., 0, 1, 2, 10) to equidistant systems (e.g., 0, 3⅓, 6⅔, 10). The monotonic scoring systems produce largely similar point-biserial coefficients, t-tests, and critical ratios. Furthermore, the divergent scoring systems are highly interrelated. The *r*'s between the two types of therapy and the four subjective responses are somewhat higher (averaging about .20) than the correlation coefficients in this study (averaging about .12).

[2] Small error may result because the difference between two adjacent ranks may not be the same as the difference between two other adjacent ranks.

To determine the degree of error of results when treating ordinal variables as if they are interval, the relation between occupational prestige and suicide rates is analyzed. Prestige rankings obtained by NORC in its 1947 survey are related to suicides by

TABLE 1. PRESTIGE, INCOME, EDUCATION, AND SUICIDE RATES FOR 36 OCCUPATIONS, UNITED STATES, MALES, *Circa,* 1950

Occupation	NORC Prestige Rating Scale[a]	Male Suicide Rate[b]	Median Income[c]	Median School Yrs. Completed[d]
Accountants and auditors	82	23.8	3,977	14.4
Architects	90	37.5	5,509	16+
Authors, editors and reporters	76	37.0	4,303	15.6
Chemists	90	20.7	4,091	16+
Clergymen	87	10.6	2,410	16+
College presidents, professors and instructors (n.e.c.)	93	14.2	4,366	16+
Dentists	90	45.6	6,448	16+
Engineers, civil	88	31.9	4,590	16+
Lawyers and judges	89	24.3	6,284	16+
Physicians and surgeons	97	31.9	8,302	16+
Social welfare, recreation and group workers	59	16.0	3,176	15.8
Teachers (n.e.c.)	73	16.8	3,465	16+
Managers, officials and proprietors (n.e.c.)—self-employed—manufacturing	81	64.8	4,700	12.2
Managers, officials and proprietors (n.e.c.)—self-employed—wholesale and retail trade	45	47.3	3,806	11.6
Bookkeepers	39	21.9	2,828	12.7
Mail-carriers	34	16.5	3,480	12.2
Insurance agents and brokers	41	32.4	3,771	12.7
Salesmen and sales clerks (n.e.c.), retail trade	16	24.1	2,543	12.1
Carpenters	33	32.7	2,450	8.7
Electricians	53	30.8	3,447	11.1
Locomotive engineers	67	34.2	4,648	8.8
Machinists and job setters, metal	57	34.5	3,303	9.6
Mechanics and repairmen, automobile	26	24.4	2,693	9.4
Plumbers and pipe fitters	29	29.4	3,353	9.3
Attendents, auto service and parking	10	14.4	1,898	10.3
Mine operatives and laborers (n.e.c.)	15	41.7	2,410	8.2
Motormen, street, subway, and elevated railway	19	19.2	3,424	9.2
Taxicab-drivers and chauffeurs	10	24.9	2,213	8.9
Truck and tractor drivers, deliverymen and routemen	13	17.9	2,590	9.6
Operatives and kindred workers, (n.e.c.), machinery, except electrical	24	15.7	2,915	9.6
Barbers, beauticians and manicurists	20	36.0	2,357	8.8
Waiters, bartenders and counter and fountain workers	7	24.4	1,942	9.8
Cooks, except private household	16	42.2	2,249	8.7
Guards and watchmen	11	38.2	2,551	8.5
Janitors, sextons and porters	8	20.3	1,866	8.2
Policemen, detectives, sheriffs, bailiffs, marshals and constables	41	47.6	2,866	10.6

[a] Albert J. Reiss, Jr., *et al.,* 1961:122–123. The scale is based on a 1947 survey.

[b] Males, aged 20–64. National Office of Vital Statistics, *Vital Statistics—Special Report,* Vol. 53, No. 3 (September, 1963).

[c] 1949 Median income. *United States Census of Population, 1950. Occupational Characteristics* (Special Report, P-E No. 1B), Table 19.

[d] 1950 Median school years completed. *Ibid.,* Table 10.

occupation for males in the United States in 1950. The list of occupations, taken from Duncan's comparisons of occupational categories used in the survey, are matched to the detailed occupational classification in the U.S. Census of 1950 (Reiss *et al.,* 1961). Because suicides are not reported for all of these occupations and sometimes the reported suicides are for two or three occupations grouped into one, 36 occupations were selected which contain the necessary data used in this study (see Table 1). Measurement of occupational prestige is based solely on the principle of ordinal ranking. In the survey, respondents were given occupations to rank by the method of paired comparisons; consequently, the resulting prestige scores indicate merely the rank of one occupation relative to the others (Reiss *et al.,* 1961:122–123).[3]

The rank correlation (*rho*) between occupational prestige and suicide is .07. The scatter diagram of the NORC prestige ratings and suicide rates suggests that the relation is roughly linear, although the plotted points are widely scattered. The Pearsonian correlation coefficient (r) on the same data is slightly larger (.11). The .04 discrepancy between the two measures is due to the magnitude of the differences between adjacent scores which are not considered in *rho,* but do influence the value of r.

ASSIGNMENT OF SCORING SYSTEMS TO ORDINAL CATEGORIES

Twenty scoring systems are used on NORC's occupational prestige values. One scoring system is the actual prestige ratings resulting from the study (the NORC Prestige Rating Scale in Table 1). A second scoring system is the assignment of equidistant numbers (i.e., an equal distance between assigned numbers) to the occupational categories (Table 2). The remaining scoring systems in Table 2 were generated from a computer according to the following conditions: (1) the assigned numbers lie between the range of 1 and 10,000, (2) the assignment of numbers is consistent with the monotonic function of the ordinal rankings, (3) any ties in the ordinal rankings are assigned identical numbers, and (4) the selection of a number is made on the basis of a random generator in the computer program. To be consistent with the monotonic function, any subsequent randomly selected numbers must be higher than previous ones (except for ties). The resulting largely random scoring systems vary among themselves (sometimes to a large extent) on the actual values assigned to each rank, the range of values, and the size of the differences between adjacent values. Although all are necessarily consistent with the monotonicity of the ordinal rankings, they vary widely among themselves. In fact, some of the scoring systems show definite curvilinear patterns—logarithmic, exponential or higher order curves (two or more inflection points).

Because this computer approach to assigning numbers to rank order data partially is based on a random selection of numbers, the generality of the findings is somewhat limited. It is possible that some systematic selection of numbers will not yield such consistent results as those reported herein.

The similarity among the scoring systems can be assessed by their matrix of intercorrelations (Table 3). By assuming, in turn, that each scoring system is the "true" one, the intercorrelations (Pearson product-moment coefficients) indicate the extent of "error" of using one of the other 19 scoring systems. For example, if (4) is the "true" system and (7) has been used in its place, then .97 (the correlation between the two scoring systems) indicates the degree to which the two systems vary together. On the other hand, r^2 (the values below the diagonal in Table 3) indicates "error" in terms of the amount of variance in the assigned scoring system accounted for by the variation in the "true" scoring system (Abelson and Tukey, 1959).[4] In this instance,

[3] Duncan's socioeconomic index, based upon the income and educational levels of each occupation, correlates highly with the NORC prestige scale.

[4] Abelson and Tukey also use r^2 as the criterion for assessing the adequacy of numerical assignments and, in addition, present a "maximin" r^2 to assess the largest possible error in a scoring system. Briefly, an assigned scoring system X is correlated with a "true" system Y so that the minimum possible r^2 between X and Y achieves its maximum value. Their analysis, instead of leading to an average error rate (in which the "true" r^2 may be equally above or below the rate), results in a conservative lower limit estimate. This lower limit

TABLE 2. NORC PRESTIGE RATINGS, LINEAR SCORING, AND FIVE MONOTONIC RANDOM GENERATED SCORING SYSTEMS[a]

Linear	NORC	Monotonic Random Generated Scoring Systems				
(1)	(2)	(3)	(5)	(9)	(13)	(18)
1.0	7	13	79	52	849	418
2.0	8	34	105	109	909	585
3.5	10	99	233	380	923	648
3.5	10	99	233	380	923	648
5.0	11	248	389	518	1152	820
6.0	13	407	580	557	1167	869
7.0	15	727	605	799	2300	1271
8.5	16	1824	771	2167	2343	1478
8.5	16	1824	771	2167	2343	1478
10.0	19	1897	1042	2790	2845	1647
11.0	20	2021	1287	2796	2876	1789
12.0	24	2470	1374	3209	3107	2112
13.0	26	2978	1713	3558	3159	2627
14.0	29	2995	2083	3598	3231	2628
15.0	33	3330	2595	3808	3409	2777
16.0	34	3412	2715	3945	3760	2921
17.0	39	3535	2751	4087	4238	3077
18.5	41	3952	2861	4094	4898	3156
18.5	41	3952	2861	4094	4898	3156
20.0	45	4082	3003	4745	5336	3209
21.0	53	4485	3266	4885	5903	3600
22.0	57	4865	4013	4892	6016	4304
23.0	59	5091	4267	5044	6106	4323
24.0	67	5146	4449	5300	6242	4762
25.0	73	5349	5318	5819	6270	5020
26.0	76	5775	6330	5876	6681	5528
27.0	81	5995	6547	5923	6787	5797
28.0	82	6304	6810	5932	6915	6027
29.0	87	6356	6974	5976	7118	6388
30.0	88	6644	7660	5995	7229	6471
31.0	89	6742	8145	6160	7652	6560
33.0	90	7657	9085	6231	7926	6911
33.0	90	7657	9085	6231	7926	6911
33.0	90	7657	9085	6231	7926	6911
35.0	93	7841	9108	6458	8283	6972
36.0	97	8164	9461	7094	8472	7588

[a] See text for an explanation of the scoring systems. The five random scoring systems are indicative of the eighteen used in the study.

between scoring systems (4) and (7), 94% of the variance in (7) is accounted for by the variation in (4).

The r and r^2 values in Table 3 are consistently and substantially high, indicating a high degree of interchangeability among the 20 scoring systems. Out of 190 correlation coefficients, all are above .90 (a few even reach unity), and 157 are .97 and above. Therefore, even without a rationale concerning the differences between ranks, by using a nearly random method of assigning scoring systems (consistent with the monotonic function), it is possible that under specific conditions the selected scoring

estimate is based on a sequence called "corners," which is consistent with the inequalities (i.e., it follows the monotonic or equality functions) and is based on a set of dichotomized values. For example, given the following relations $Y_1 \leq Y_2 \leq Y_3 \leq Y_4$, a set of corners is (0, 0, 0, 1), (0, 0, 1, 1), and (0, 1, 1, 1). One of these corner sequences yields the maximin r^2. There are three problems with Abelson and Tukey's analysis: (1) an average error rate is more indicative of a representative error (i.e., the most likely error in assigning a scoring system) and, therefore, is more useful to the researcher, (2) the corner sequence is based on dichotomies which is a highly unlikely occurrence and a waste of information, and (3) they analyze only "greater than" *and* "equal to" models in combination ($Y_3 \geq Y_2 \geq Y_1$), while the most frequent ordinal cases are "greater than" between most ranks ($Y_3 > Y_2 > Y_1$). The "greater than" model leads into a dichotomous analysis only if there are two ranks (a trivial case).

TABLE 3. INTERCORRELATIONS (r) AMONG TWENTY SCORING SYSTEMS[a]

Scoring Systems	Scoring Systems (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1)[b]	...	.98	1.00	.99	.97	.99	.98	.99	.97	.98	.99	.99	.99	.99	1.00	.98	.99	.99	.99	1.00
(2)[c]	.98	...	.97	.96	.97	.98	.95	.97	.96	.94	.97	.97	.96	.98	.93	.97	.94	.99	.98	.96
(3)	1.00	.94	...	.99	.97	.99	.98	.98	.98	.98	.99	.99	.99	.99	.99	.98	.99	.99	.99	.99
(4)	.98	.92	.98	...	.98	.99	.97	.99	.95	.99	.98	.98	.98	.99	.99	.99	.99	1.00	.98	.99
(5)	.94	.94	.94	.96	...	.99	.93	.99	.91	.99	.96	.95	.96	.96	.97	.99	.98	.99	.94	.96
(6)	.98	.96	.98	.98	.98	...	.97	.99	.94	.99	.97	.98	.98	.98	.99	.99	.99	.99	.97	.98
(7)	.96	.90	.96	.94	.86	.94	...	.96	.98	.95	.96	.98	.98	.98	.98	.94	.96	.96	.99	.98
(8)	.98	.94	.96	.98	.98	.98	.92	...	.93	.99	.97	.97	.98	.98	.99	.99	.98	.99	.96	.98
(9)	.94	.92	.96	.90	.83	.88	.96	.86	...	.93	.97	.98	.97	.97	.96	.92	.96	.95	.99	.98
(10)	.96	.88	.96	.98	.98	.98	.90	.98	.86	...	.97	.97	.97	.98	.98	.99	.99	1.00	.96	.98
(11)	.98	.94	.98	.96	.92	.94	.92	.94	.94	.94	...	.99	.98	.98	.99	.98	.99	.98	.98	.99
(12)	.98	.94	.98	.96	.90	.96	.96	.94	.96	.94	.98	...	.99	.99	.98	.97	.98	.98	.99	.99
(13)	.98	.92	.98	.96	.92	.96	.96	.96	.94	.94	.96	.98	...	.99	.99	.97	.98	.98	.99	.99
(14)	.98	.96	.98	.98	.92	.96	.96	.96	.94	.96	.96	.98	.98	...	.99	.97	.98	.99	.98	.99
(15)	1.00	.86	.98	.98	.94	.98	.96	.98	.92	.96	.98	.96	.98	.98	...	.98	.99	.99	.98	.99
(16)	.96	.94	.96	.98	.98	.98	.88	.98	.85	.98	.96	.94	.94	.94	.96	...	.99	.99	.96	.97
(17)	.98	.88	.98	.98	.96	.98	.92	.96	.92	.98	.98	.96	.96	.96	.98	.98	...	.99	.98	.99
(18)	.98	.98	.98	1.00	.98	.98	.92	.98	.90	1.00	.96	.96	.96	.98	.98	.98	.98	...	.97	.98
(19)	.98	.96	.98	.96	.88	.94	.98	.92	.98	.92	.96	.98	.98	.96	.96	.92	.96	.94	...	.99
(20)	1.00	.92	.98	.98	.92	.96	.96	.96	.96	.96	.98	.98	.98	.98	.98	.94	.98	.96	.98	...

[a] r above the diagonal; r^2 below.
[b] linear scoring system.
[c] NORC prestige ratings.

system will deviate from the "true" system by a near zero or negligible amount. The r^2 values are slightly lower than the r values, but still exceedingly high. For example, only nine of the 190 are below .90, and none are below .83. (Since r^2 is the square of a decimal fraction, it is necessarily smaller than r.)

Note that if the equidistant (linear) scoring system is always selected (no matter what the "true" scoring system may be), the expected error is smaller than the larger errors cited above. Almost all the r's and r^2's for the linear system (1) are near unity, with the lowest r being .97 and the lowest r^2 being .94. The linear scoring system lies midway between the other scoring systems (in correlational terms), which by definition excludes the most extreme scoring systems in each direction. The correlations between the extremes are lowest, and, therefore, selecting the linear scoring system eliminates the lowest r's and the highest potential "errors" in selecting a scoring system different from the "true" one.

Possessing some knowledge about the amount of differences between ranks can reduce the small error even further, if the linear scoring system has been assigned to the ordinal categories. Perhaps, the best strategy, if there is some knowledge of the differences between ranks, is to modify the linear scoring system accordingly. For example, in the relation $X_1 > X_2 > X_3$, X_2 is assumed to be closer to X_3 than to X_1. Consequently, the linear scoring system of 10, 20 and 30 (as values for X_1, X_2 and X_3) can be modified to 10, 25 and 30 to account for this additional knowledge. It should be stressed that without prior knowledge or theory such score assignments are not likely to prove useful for analysis.

Table 4 offers further evidence that ordinal data can be treated as if they are interval by assigning scoring systems to the ordered categories. In this instance, the predictive ability of each scoring system is assessed in terms of its relation to suicide rates. As indicated previously, the *rho* value between the NORC prestige scale and 1950 suicide rates for males in 36 occupations is .07; for the same data, r is .11. Table 4 reports the r and r^2 values between the 20 scoring systems and the suicide rate. (The last two columns in Table 4 are r values for 20 and 10 occupations respectively and will be discussed later in the paper.) The similarity in predicting an outside variable is extremely high. The r's vary between .09 and .15, and the r^2 values are either .01 or .02.[5] Given some degree of unreliability in

TABLE 4. CORRELATION COEFFICIENTS (r) BETWEEN SUICIDE RATES AND TWENTY SCORING SYSTEMS OF OCCUPATIONAL PRESTIGE[a]

Scoring System	r(N=36)	r^2	r(N=20)	r(N=10)
(1) (linear)	.13	.02	.35	.28
(2) (prestige ratings)	.11	.01	.35	.25
(3)	.13	.02	.31	.24
(4)	.11	.01	.32	.30
(5)	.10	.01	.30	.21
(6)	.11	.01	.35	.18
(7)	.14	.02	.28	.33
(8)	.12	.01	.38	.34
(9)	.14	.02	.26	.15
(10)	.09	.01	.29	.24
(11)	.13	.02	.30	.24
(12)	.11	.01	.28	.22
(13)	.15	.02	.41	.35
(14)	.14	.02	.37	.25
(15)	.13	.02	.35	.32
(16)	.09	.01	.33	.18
(17)	.12	.01	.37	.34
(18)	.11	.01	.30	.33
(19)	.15	.02	.33	.25
(20)	.14	.02	.38	.41

[a] Partially based on the data in Tables 1 and 2. Scoring systems 3–18 are randomly generated.

[5] It should be noted that the usual purpose of a transformation in correlation work is to raise the

occupational prestige and suicide data, and the rather crude measurement procedures, these results substantiate the point that different systems yield interchangeable variables. Each indicates a quite low positive (statistically nonsignificant) relation between occupational prestige and suicide. These results are consistent with a previous study (Labovitz, 1967), which also found the relations to be very similar; however, in the previous study, the relations are somewhat higher and statistically significant.

As indicated by the results in Table 4, the greater the number of ranks (N), the greater the confidence in assigning an interval scoring system to ordinal data (Labovitz, 1968a; Morris, 1968). The last two columns (Table 4) report the correlations for the first twenty and for the first ten occupational groups (between suicide rates and the 20 scoring systems). That the correlation coefficients based on smaller N's are appreciably higher than for $N = 36$ is not a major concern. A statistical explanation for the higher *r*'s with smaller N's is that in this case and for whatever reason, the error variance diminishes more rapidly than the total variance as N increases. By restricting the range in a systematic manner, i.e., taking the first twenty and first ten, some occupations with rather high suicide rates and low prestige levels are eliminated. The net effect is an increase in the positive correlation. The major concern is not with the magnitudes but with the similarity of coefficients within each of the three groups ($N = 10$, $N = 20$, and $N = 36$). The standard deviations among the correlation coefficients decrease as N increases: (1) for $N = 10$; $SD = .07$; (2) for $N = 20$; $SD = .04$; and (3) for $N = 36$; $SD = .02$.

Note the similarity between the equidistant correlation coefficient (scoring system 1) and the mean correlation coefficient for all 20 scoring systems for each of the three groups. For 10 occupations, the mean correlation is .27 and the equidistant correlation is .25; for twenty occupations, the mean is .33 and the equidistant is .35; and finally, for all 36 occupations, the mean is .12 and the equidistant is .11. This lends further support to the suggested strategy of imposing an equidistant scoring system on ordinal categories.

It may be argued that these results do not hold for "extreme" nonlinear monotonic transformations of ordinal measures. Admittedly, there is a point beyond which the transformation will not yield interchangeable measures. This is the case where the assigned scoring system has essentially dichotomized the ranks. For example, if the true or assigned system is 1, 2, 3, 4, and the assigned or true system is 1, 9,996, 9,998, 10,000, then we are in essence scoring the categories as 1, 10,000, 10,000, 10,000. Under such conditions, it is obvious that treating ordinal categories as if they are interval is not an aid in data analysis, unless the "dictohomy" is recognized. The problem of dichotomizing becomes increasingly serious (in terms of faulty interpretations) as the number of ranks increases. Under the condition of a true equidistant scoring system and exponential power function or logarithmic power function assigned scoring systems, as k increases the true and assigned systems increasingly diverge in values.

To illustrate this last point, consider an equidistant scoring system for X, e.g., 1, 2, 3, 4. Suppose X' (assigned scoring systems) is set equal to X^k and scores are generated. For given values of k the scoring systems are: (1) 1, 2, 3, 4 ($k = 1$); (2) 1, 4, 9, 16 ($k = 2$); (3) 1, 8, 27, 64 ($k = 3$); (4) 1, 16, 81, 256 ($k = 4$); and (5) 1, 32, 243, 1024 ($k = 5$). As noted above, as k increases the new scoring systems come increasingly closer to dichotomizing (polarizing) the values. The equidistance between numbers is progressively lost; the fourth number becomes large more rapidly than the first, second, or third; the third number increases more rapidly than the second; and the second more rapidly than the first. Consequently, the most deviant result in this process (i.e., the farthest from the true scoring system) is polarization into an essential dichotomy for very large k's.

Although we generally can partition a variable into more than two intervals, it is useful to consider the correlation between the linear system (1, 2, 3, 4) and the di-

correlation. However, the stability of the correlations in this study is not inconsistent with this general principle.

chotomies (0, 0, 0, 1), (0, 0, 1, 1), (1, 1, 0, 0), and (1, 0, 0, 0). The correlations are respectively .77, .83, −.83, and −.77. The correlation of .77 (or −.77) represents the lower limit (the worst possible situation) for the four number sets compared to the equidistant scoring system (1, 2, 3, 4). With regard to an "outside" variable, "extreme" transformations yielding a dichotomy do make a difference. The equidistant system X (1, 2, 3, 4) and the exponential system X′ (1, 8, 27, 64) are related to Y (2, 1, 3, 5) as follows: $r_{xy} = .83$ and $r_{x'y} = .94$. In this case, X and X′ are not interchangeable variables.

Perhaps the most important reason for treating an ordinal variable as if it conforms to an interval scale lies in the opportunity it provides for applying well-developed and interpretable multivariate techniques. Although partials can be applied to ordinal measures—e.g., partial *tau*, partial *gamma*, or partial *rho*—these are often difficult to interpret. Further, a multiple relationship measure is not defined for ordinal variables, unless they are assumed to be the counterpart of the correlation coefficient.

To illustrate the utility of using multivariate analysis, education and income are combined with the equidistantly (linearly) scored prestige scale to account for the variance in suicide rates. Although occupational prestige is not highly related to suicide (in a zero-order correlation), when combined with an additive combination of income, the multiple *R* is .31. The zero-order correlation between income and suicide is .26. An additive combination of occupational prestige, income, and education results in an *R* of .55. This represents an increase of .05 over the *R* of .50 between suicide and the independent variables of income and education. Over 30% of the variance in suicide rates is accounted for by an additive combination of the three major independent variables. Treating occupational prestige as ordinal would not have permitted this analysis, although this variable adds the least amount of explained variance to the multiple *R*.

Partial correlations also result in significant findings. When partialed on income, the relation between occupational prestige and suicide is negative (−.16). Although the relation between the two is still quite low, a reasonable interpretation of the partial correlation is that income is determining the positive association between the two (by its positive effect on both variables). An implication is that conflicting results among several studies between prestige and suicide may be resolved by controlling for variations in income (Powell, 1958; Hirsh, 1959; Dublin, 1963; Breed, 1963; Labovitz, 1968b; Maris, 1967).

Consistent with the above partial of −.16, the relation between income and suicide increases when the effects of occupational prestige are partialed out. It appears that occupational prestige acts as a suppressor variable in relating income to suicide, and, therefore, since prestige and income are differentially related to suicide (one positive and one negative), they are to some extent canceling out each other's effects.

With regard to multivariate analysis, the treatment of occupational prestige as an interval variable has several advantages. First, a rather small N (in this instance the number of occupations is only 36) can be used with most intervally based multivariate techniques. Partialing by modes of elaboration techniques (cross-tabulation) may require an extremely large N. Second, these techniques are well-developed for interval data, but are either not developed or poorly developed for ordinal data. Consequently, the degree of latitude or versatility (Anderson, 1961) [6] in statistical analysis is increased substantially by using interval statistics. Finally, and in summary, the multivariate analysis led to some highly suggestive conclusions that would have been overlooked if prestige was treated as an ordinal variable. For example, (1) income may be determining the positive relation between prestige and suicide, (2) prestige may be suppressing or depressing the relation between income and suicide, and (3) the predictive model of an additive combination of prestige, income, and education accounts for a moderate part (30%) of the variance in suicide.

The researcher should be warned that the similarity among the scoring systems in

[6] This is Anderson's basic reason for selecting parametric over nonparametric statistics.

terms of intercorrelations and predictive ability should not be uncritically generalized to regression problems. The bivariate regression coefficients (slopes) of the scoring systems on suicide rates for 36 occupations range from .04 to 32.8. This substantial variation is largely due to the linear and NORC scoring systems which used smaller numbers than those generated by the computer. The computer generated systems are closer in slope values, ranging from 18.2 to 32.8. The range of the standardized zero-order (gross) regression coefficients stand in sharp contrast to these results. The bivariate standardized slopes (which are the zero-order correlation coefficients in the standard score regression line) range from the lowest of .09 to the highest of .15.

These results indicate that if scores are assigned to an ordinal system (or if there is uncertainty regarding the magnitude of the differences between adjacent scores), regression coefficients should be standardized because of their greater stability from one scoring system to another. Standardized coefficients are an integral part of path analysis and are sometimes used as the path coefficients. Interpretations of path coefficients (standardized), according to the findings given above, do not appear to require modification by the assignment of numbers to ordinal categories. A wide range of values would indicate unreliable coefficients and would negate any meaningful solution of the identification problem (that is, estimating the unknown parameters in a model from available empirical data). However, as Blalock (1967 and 1968) has pointed out, standardized as compared to unstandardized coefficients may not be as adequate for problems where the comparison of populations is necessary to determine whether or not the underlying causal processes are basically similar. Standardized coefficients appear to be more adequate for problems of generalizing to a specific population, because they can be used to assess the direct and joint contributions of the several independent variables.

CONCLUSIONS

The results of the tests based on assigning interval scores to ordinal categories suggest: (1) certain interval statistics can be used interchangeably with ordinal statistics and interpreted as ordinal, (2) certain interval statistics (e.g., variance) can be computed where no ordinal equivalent exists and can be interpreted with accuracy, (3) certain interval statistics can be given their interval interpretation with only negligible error if the variable is "nearly" interval, and (4) certain interval statistics can be given their interval interpretations with caution (even if the variable is "purely" ordinal), because the "true" scoring system and the assigned scoring system, especially the equidistant system, are almost always close as measured by r and r^2.

Consequently, treating ordinal variables as if they are interval has these advantages: (1) the use of more powerful, sensitive, better developed and interpretable statistics with known sampling error, (2) the retention of more knowledge about the characteristics of the data, and (3) greater versatility in statistical manipulation, e.g., partial and multiple correlation and regression, analysis of variance and covariance, and most pictorial presentations.

The study suggests two research strategies when analyzing ordinal variables. First, assign a linear scoring system according to the available evidence on the distances between ranks. Second, use all available rank order categories, rather than collapsing them into a smaller number, because the greater the number of ranks the greater the stability and confidence in the assigned scoring system (unless the dichotomization of ranks is suspected). The all-too-frequent strategy of dichotomizing or trichotomizing variables should be avoided if possible.[7]

A final word of caution is necessary. The researcher should know and report the actual scales of his data, and any interval statistics selected should be interpreted with care. Further exploration and tests are necessary for added confidence in treating ordinal data as if they are interval. The more conservative procedure, of course, is to treat ordinal data as strictly ordinal, and thereby avoid the possibility of attributing

[7] Another reason against the use of dichotomies or trichotomies is that often a large amount of information is lost by such drastic collapsing.

a property to a given scale which it does not possess.

REFERENCES

Abelson, Robert P. and John W. Tukey
1959 "Efficient conversation of non-metric information to metric information." Proceedings of the Social Statistics Section, American Statistical Association: 226–230.

Anderson, Norman A.
1961 "Scales and statistics: Parametric and nonparametric." Psychological Bulletin, 58 (July):305–316.

Blalock, Hubert M., Jr.
1967 Letter. American Journal of Sociology 72 (May):675–677.
1968 "Theory building and causal inferences." Pp. 155–196 in Hubert M. Blalock, Jr. and Ann B. Blalock (eds.), Methodology in Social Research. New York: McGraw-Hill.

Boneau, C. A.
1960 "The effects of violations of assumptions underlying the 't' test." Psychological Bulletin, 57 (January):49–64.

Breed, Warren
1963 "Occupational mobility and suicide among white males." American Sociological Review, 28 (April):179–189.

Dublin, L. I.
1963 Suicide: A Sociological and Statistical Study. New York: Ronald Press.

Hirsh, Joseph.
1959 "Suicide." Mental Hygiene, 43 (October): 516–526.

Labovitz, Sanford
1967 "Some observations on measurement and statistics." Social Forces, 46 (December): 151–160.
1968a "Reply to Champion and Morris." Social Forces, 46 (June):543–545.
1968b "Variation in suicide rates." Pp. 57–74 in Jack P. Gibbs (ed.), Suicide. New York: Harper and Row.

Lindquist, E. F.
1953 Design and Analysis of Experiments. Pp. 78–90. New York: Houghton Mifflin.

Maris, Ronald
1967 "Suicide, status, and mobility in Chicago." Social Forces, 46 (December):246–256.

Morris, Raymond N.
1968 "Commentary." Social Forces, 46 (June): 541–543.

Powell, Elwin M.
1958 "Occupation, status, and suicide: Toward a redefinition of anomie." American Sociological Review, 23 (April):131–140.

Reiss, Albert J., Jr., with O. D. Duncan, Paul K. Hatt, and Cecil C. North
1961 Occupations and Social Status. New York: Free Press.

Somers, Robert H.
1962 "A new asymmetric measure of association for ordinal variables." American Sociological Review, 27 (December):799–811.