PEARSON'S R AND COARSELY CATEGORIZED MEASURES*

KENNETH A. BOLLEN
KENNEY H. BARB
General Motors Research Laboratories

Errors can be introduced into scientific research when continuous concepts are measured on scales that rank the concepts into a few categories. This presents a potential problem because measures of association between two variables may differ depending on whether continuous or collapsed measures are used. We analyzed simulated data and examined differences in the correlation between two normally distributed continuous variables and the same two variables collapsed into a small number of categories. In general, the differences in correlation coefficients computed on continuous variables and the same variables collapsed into a few categories are small.

The greatest differences in the correlations between the two types of variable occur when the continuous variables' correlation is high and only a few categories are used for the collapsed variables. When as few as five categories are used to approximate the continuous variables, the correlation coefficients and their standard deviations for the collapsed and continuous variables are very close. These findings suggest that under certain conditions it may be justifiable to analyze categorical data as if it were continuous.

Measurement problems continue to be one of the main obstacles to the advancement of social science research (Blalock, 1979). One manifestation of this problem is the measurement imprecision resulting when continuous concepts are measured on scales containing relatively few categories. Attitudes, for instance, generally are measured on Likert scales consisting of only five categories even though the underlying attitudes may be continuous. Even less abstract concepts such as income and age frequently are reported in broad categories rather than in their continuous form. Consequently, researchers must often estimate relationships between continuous concepts with categorized measures. This is problematic because measures of association can differ depending on whether continuous or categorical variables are used. Furthermore, when bivariate correlation coefficients based on categorized measures are used in more complex forms of multivariate analysis (e.g., multiple regression, factor analysis) their inaccuracies may distort the subsequent analyses. Therefore, it is important to examine the size of the differences between correlations computed with categorized and continuous variables to determine the seriousness of measurement imprecision.

The general problems of correlating categorized, grouped, or collapsed versions of continuous measures have been examined by statisticians (Lancaster and Hamdan, 1964; Richie-Scott, 1918; Pearson, 1913, 1904; Pearson and Heron, 1913), market researchers and psychologists (Lehmann and Hulbert, 1972; Green and Rao, 1971, 1970; Benson, 1971; Jacoby and Matell, 1971; Komorita and Graham, 1965; Bendig, 1954; Symonds, 1924). Most sociological research has focused on the "interval-ordinal debate" (e.g., Kim, 1978, 1975; Grether, 1976; Labovitz, 1975, 1972, 1971, 1970, 1968, 1967; Henkel, 1975; Acock and Martin, 1974; Mayer, 1971, 1970; Schweitzer and Schweitzer, 1971; Vargo, 1971; Wilson,

* Direct all communications to: Kenneth A. Bollen, Societal Analysis Department, General Motors Research Laboratories, Warren, Michigan 48090.

We wish to thank several members of the Societal Analysis Department for their contribution to this paper: Dana B. Kamerud and Donald E. Schmidt provided helpful substantive and editorial comments. Thomas J. Laffey assisted with the computer programming, and Linda E. Meyer typed several drafts as well as the final version of this paper.

1971; Champion, 1968; Morris, 1968; Baker, et al., 1966). One side argues that ordinal scores, which represent some monotonic transformation of the true scores, should only be analyzed with "ordinal level" statistics—those statistical techniques that do not take into account the distance between scale points. The other side argues that it is better to treat data that are theoretically interval as if they were interval, even though the measurements may contain only ordinal information. They contend that parametric or "strong statistics" can address more theoretically interesting questions, are more amenable to multivariate analyses, and are robust enough to be only slightly distorted by most monotonic transformations. Some advocates of "strong statistics" have attempted to develop ways to assign scores to the ranked data so as to provide some information about the distance between rankings (Hensler and Stipak, 1979; Allen, 1976).

Labovitz has been the most influential advocate for treating ordinal data as interval. He demonstrated that certain types of interval statistics were reasonably accurate even when the distance between scale points was randomly distorted. A recent research note by O'Brien (1979) expands Labovitz' work. O'Brien found that, contrary to previous suggestions in the literature, Pearson's r between an interval variable and its categorized version was not a monotonic function of the number of categories used to rank the data. In fact, in the case of a skewed distribution, r was a decreasing function of the number of ranks. O'Brien's work focuses on how the collapsed or ranked variable correlates with its "true" continuous or interval measure.

It remains unclear, however, if O'Brien's results generalize to the bivariate case. For example, he finds that the correlation between a collapsed variable and its continuous form is a nonmonotonic function of the number of categories. Does this mean that the bivariate correlation between two collapsed variables will be a nonmonotonic function of the number of categories? O'Brien (1979:855) also concluded that "if one is willing to assume that the underlying variable is either uniformly or normally distributed, the use of rank-order values causes little distortion in r." Does this conclusion also apply when two different rank-ordered (or categorized) variables are correlated? After all, most researchers are concerned with the correlation between *different* categorized variables rather than the correlation of a variable with a collapsed version of itself.

This research note explores two questions. First, how is the correlation coefficient between continuous concepts affected by analyzing collapsed scales rather than continuous scales? Second, are the number of categories necessary for accurate results the same with differing correlations between the continuous variables?

DESCRIPTION OF SIMULATION

Two factors are varied in our simulation: the number of categories into which the variables are collapsed, and the strength of the relationship between the continuous variables. First, each normally distributed variable in the simulation is collapsed into a number of categories ranging from two through ten (see Figure 1). This range of categories includes many of the attitudinal scales and other "rank" variables commonly encountered in research.

The original, normally distributed continuous variable is shown at the top of

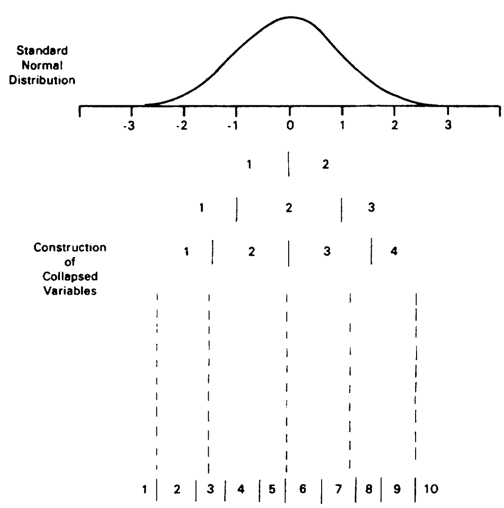

Figure 1. Construction of Collapsed Variables from Standard Normal Distribution

Figure 1. The cutting points which form the new collapsed variable are beneath the graph of the standard normal distribution. The cutting points are determined by considering the area under the normal curve between ±3 standard deviations. On the average this range includes nearly all (i.e., over 99.7%) of the cases. The six-unit interval between ±3 is divided by the number of categories to determine the number of unit intervals per category.

Consider the three-category case. The six intervals are divided by three, which leads to two intervals for each category. So if the value of the normally distributed variable is less than or equal to −1, the collapsed variable is coded as 1. If the normal variable is greater than −1 and less than or equal to +1, the collapsed variable is coded 2. Finally, those values of the normal variable greater than 1 are coded as 3.

Collapsing in this manner leads to variables that are symmetric and, as the number of categories increases, to variables that approximate a normal distribution. Other distributions of the collapsed variable could be created, but as a first step we assumed the collapsed variable has a distribution similar to the underlying continuous variable.

The second factor varied in the simulations is the strength of the relationship between the original continuous variables. Each pair of normally distributed variables is constructed to correlate at one of five magnitudes: 0.2, 0.4, 0.6, 0.8, or 0.9. Fifty samples of 500 observations are generated for each correlation. The variables are then collapsed into a smaller number of categories and the resulting correlation between the collapsed variables (henceforth "collapsed correlation") is compared to that of the original continuous variables ("continuous correlation").

Table 1. Comparison of Mean Correlation Coefficients for Original and Collapsed Variables

Number of Categories	Mean Collapsed Correlation				
2	0.117	0.264	0.414	0.579	0.719
3	0.149	0.302	0.445	0.617	0.727
4	0.165	0.336	0.504	0.669	0.774
5	0.182	0.359	0.533	0.712	0.806
6	0.180	0.371	0.554	0.733	0.833
7	0.192	0.380	0.563	0.750	0.850
8	0.193	0.384	0.571	0.760	0.860
9	0.194	0.389	0.577	0.771	0.870
10	0.197	0.391	0.581	0.772	0.875
Mean Original Correlation	0.203	0.404	0.599	0.796	0.901

Based on 50 samples of 500 observations each.

RESULTS

Table 1 presents the average continuous correlation (last row) and the average collapsed correlations. All the collapsed correlations are less than the continuous correlations. The greater the number of categories, however, the closer the collapsed correlations are to those of the original continuous correlations. In the first column of Table 1, for instance, where the continuous r = 0.203, the collapsed correlations change from a low of 0.117 for the two-category coding to a high of 0.197 when ten categories are used. The ten-category coding produces a collapsed correlation nearly the same as that of the continuous correlation.

In virtually all cases the differences between the continuous correlation and the collapsed correlations are small. The largest differences of about 0.2 occur only in the upper right-hand side of Table 2, where the highest correlations are combined with the lowest number of categories. For example, the correlation of the three-category coding in the last column of Table 2 is 0.174 less than the continuous correlation of 0.901. This dif-

Table 2. Difference of Mean Correlation Coefficients of Original and Collapsed Variables

Number of Categories	Mean Differences				
2	0.086	0.140	0.185	0.217	0.182
3	0.054	0.102	0.154	0.179	0.174
4	0.038	0.068	0.095	0.127	0.127
5	0.021	0.045	0.066	0.084	0.095
6	0.023	0.033	0.045	0.063	0.068
7	0.011	0.024	0.036	0.046	0.051
8	0.010	0.020	0.028	0.036	0.041
9	0.009	0.015	0.022	0.025	0.031
10	0.006	0.013	0.018	0.024	0.026
Mean Original Correlation	0.203	0.404	0.599	0.796	0.901

ference is more than cut in half when the number of categories is increased from three to six. In fact, the difference between the average correlation for the continuous and the collapsed variables is always less than 0.1 when five or more categories are used.

The percentage of the continuous correlation reproduced by the collapsed correlation (Table 3) provides another perspective on this problem. The collapsed correlation was divided by the continuous correlation and multiplied by 100 to form a percentage. The higher the percentage the closer the collapsed correlation is to the continuous correlation. For instance, when the mean continuous correlation is 0.599, (Table 3, column 3) the mean correlation between the collapsed variables with five categories is about 90% the magnitude of the original correlation.

Table 3. Mean Correlation of Collapsed Variables as Percentage of the Mean Correlation of Original Variables

Number of Categories	Percentages				
2	57.6	65.3	69.1	72.7	79.8
3	73.4	74.8	74.3	77.5	80.7
4	81.3	83.2	84.1	84.0	85.9
5	89.7	88.9	89.0	89.4	89.5
6	88.7	91.8	92.5	92.1	92.5
7	94.6	94.1	94.0	94.2	94.3
8	95.1	95.0	95.3	95.5	95.4
9	95.6	96.3	96.3	96.9	96.6
10	97.0	96.8	97.0	97.0	97.1
Mean Original Correlation	0.203	0.404	0.599	0.796	0.901

The lowest percentage reproduction of the continuous correlation occurs when the magnitude of the relationship is low and few categories are used. For example, in the two-category coding (Table 3, row 1) where the continuous correlation increases from 0.203 to 0.901, the percentage of the correlation reproduced varies from about 58% to 80%. As the number of categories increases the percentage of the original correlation reproduced increases. For example, in the first column where the original correlation was 0.203, the two-, five-, and ten-category variables reproduce 58%, 89%, and 97% of the continuous correlation.

Differences in the percentage of the continuous correlation reproduced diminish as the number of categories increases. These differences disappear rapidly as five or more categories are used. For instance, the percentage of reproduced correlation is close to 89% for all columns in the five-category case. This convergence suggests a possible "threshold effect" when five or more categories are used. (The threshold effect is more clearly illustrated in Figure 2.) The percentage of the correlation between the continuous variables reproduced by the collapsed variables is shown on the vertical axis. The number of categories is the horizontal axis. The graph plots the percentages for three different original correlations: the lower line is for the 0.203 original correlation, the middle line for the 0.599, and the top line for the 0.901 correlation. The differences in the percentages are greatest for the two-category case. The differences rapidly converge so that with five or more categories, only 11% or less of the correlation is not reproduced, regardless of the magnitude of the original correlation.

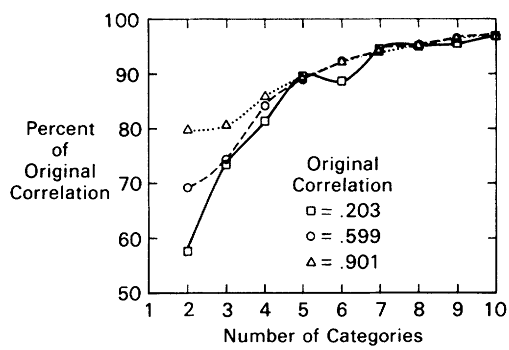

Figure 2. Percent of Original Correlation Reproduced by Number of Categories

An additional noteworthy characteristic is the comparison between the standard deviations of the correlations computed on collapsed variables and the standard deviations of the correlations of the continuous variables. Even though the average collapsed correlation is not much lower than the continuous correlation, it is possible that the average standard deviation of the collapsed correlations is greater than that of the continuous correlations. If this is true, the collapsed correlations would have greater fluctuations around their mean value than the continuous correlations. This would weaken the efficiency of using the collapsed correlations as an appropriate substitute for the continuous correlation.

To examine this issue, the standard deviation of the correlations for the collapsed variables and the original continuous variables are computed. Each combination of the number of categories and strength of relationship has 50 correlations. The bottom row of Table 4 reports the standard deviation of the correlations computed on the continuous variables; the columns above the continuous variables show the standard deviations of the correlations between the collapsed variables.

Table 4. Standard Deviations of Correlation Coefficients for the Continuous and Collapsed Variables

Number of Categories	Correlation Between Original Variables 0.203	0.404	0.599	0.796	0.901
2	0.049	0.047	0.044	0.036	0.028
3	0.040	0.043	0.030	0.032	0.027
4	0.048	0.046	0.033	0.027	0.017
5	0.042	0.041	0.030	0.022	0.011
6	0.041	0.042	0.030	0.018	0.013
7	0.040	0.038	0.033	0.021	0.010
8	0.044	0.040	0.029	0.016	0.009
9	0.042	0.040	0.031	0.019	0.008
10	0.042	0.040	0.029	0.018	0.007
Original Variables	0.042	0.040	0.030	0.017	0.007

Based on 50 samples of 500 observations each.

The largest standard deviations for the collapsed correlations, relative to the deviations of the continuous correlation, is in the upper right-hand corner of Table 4. When the original correlation is high and the number of categories is low, the standard deviations of the correlations for the collapsed variables are much larger than the standard deviations of the correlations of the original variables. However, with five or more categories the standard deviations of the correlations for the collapsed and continuous variables are close. This apparent threshold at five categories provides additional credence to the importance of using at least five, preferably more, categories in designing questionnaires.

Results with Nonnormal Distributions

We did not address the question of what happens if the underlying distribution is not normally distributed. Whether our results will still hold must await separate simulations or analytic work for a number of nonnormal distributions for a complete answer. However, some evidence exists to address this issue. An unpublished simulation study by Stern (1971) examined two uniformly distributed independent variables that were strongly related to a dependent variable. Her findings with uniformly distributed explanatory variables led to results similar to ours.[1] Wylie (1976) examined the correlations of collapsed variables that have extreme opposite skews, moderately opposite skews, and slight opposite skews. He found that the greatest differences in correlations occurred when the continuous variables were dichotomized and had extreme opposite skews. Smaller differences occurred when more than two categories were used.[2]

An actual empirical example provides additional insights on the effects of collapsing skewed variables. Even though a single example is not definitive, it does

[1] Our study differs from Stern's work in several ways. First, Stern examined two uniformly distributed independent variables, whereas we focus on two normally distributed variables. Second, her independent variables were strongly related to the dependent variable, but we vary the strength of the relationship from weak ($r = 0.203$) to strong ($r = 0.90$). Finally, Stern's original variables only had 21 categories, which were grouped into a smaller number of categories, whereas our original continuous variable had a theoretical range from $\pm\infty$.

[2]. Wylie's study differs from ours because he examined skewed distributions compared to our normal distributions. Furthermore, Wylie's original variables only had 24 possible categories, which were subsequently collapsed into a smaller number. As already discussed, our continuous variables had a theoretical range of $\pm\infty$. This latter point permits us to address what we think is a theoretically relevant point that Stern and Wylie do not address. In a sense, Stern and Wylie are asking whether one can reasonably collapse a few categories, 21 and 24, respectively, into a smaller number without too much error. We assume, however, that many variables are theoretically continuously distributed across a wider range of values, and hence we ask whether one can collapse a concept which in theory can range from $\pm\infty$ into five or six categories.

provide an indication of the difference in correlations computed on categorized versions of positively skewed continuous variables. The example is realistic because it represents what is done in practice rather than a simulation. The two variables are land area and the population of nations circa 1960. Both variables are positively skewed.[3] The continuous measures of these variables are reported in Russet et al. (1964). Coincidentally these variables are collapsed into four categories by Banks and Textor (1963).

The correlation matrix of the continuous and collapsed variables is reported in Table 5. Several results are worth noting. First, the difference between the correlation of the continuous measures (r_{XY} = 0.53) and that of the collapsed variables (r_{X*Y*} = 0.51) is only 0.02. Our simulation results in Table 2 show that when the original correlation is around 0.5 and the variables are categorized into four groups, one would expect a difference of 0.07 to 0.09 with normally distributed variables. Therefore, the 0.02 difference is less than would be expected.

Table 5. Correlation Matrix of Continuous and Collapsed Measures of Land Area and Population, Circa 1960 (N = 113)

	X	X*	Y	Y*
X	1.00			
X*	0.63	1.00		
Y	0.53	0.39	1.00	
Y*	0.48	0.51	0.56	1.00

X = continuous measure of land area
X* = collapsed measure of land area
Y = continuous measure of population
Y* = collapsed measure of population

SOURCES: Banks and Textor (1963), Russett et al. (1964).

A second interesting result is that the correlations between either of the continuous variables and the other collapsed variable are considerably less than the correlation of the two collapsed variables. Correlation of the continuous measure of land area (X) with the collapsed population variable (Y*) is 0.48, while the correlation of the collapsed measure of land area (X*) with the continuous population variable (Y) is 0.39 compared to r_{X*Y*} of 0.51. So in this example, correlating the two collapsed measures produces a correlation closer to that of the continuous measures than correlating one continuous and one collapsed variable.

Finally, the correlations of the continuous with the collapsed versions of the same variables are relatively low (r_{XX*} = 0.63, r_{YY*} = 0.56). This is particularly surprising given the closeness of r_{X*Y*} and r_{XY}. If r^2_{XX*} and r^2_{YY*} were used as estimates of reliability of X* and Y*, respectively, it would suggest that X* and Y* are very unreliable measures. Furthermore, dividing r_{X*Y*} by the square root of the product of the reliabilities r^2_{XX*} and r^2_{YY*} to correct for measurement-error attenuation would yield an unattenuated correlation of 1.5 for the correlation between X and Y! This impossible correlation may be due to nonrandom measurement error in the collapsed variables. The correction for attenuation assumes that errors of measurement for X* and Y* are uncorrelated. However, it seems likely that a positive correlation between the measurement error involved in grouping X* (population) and Y* (land area) exists. This would tend to increase the r_{X*Y*} over what it would be with independent measurement error. Thus, the correlation between the collapsed measures of population and land area (r_{X*Y*}) may be close to the correlation of the continuous measures (r_{XY}) for the "wrong reason."

DISCUSSION AND CONCLUSIONS

Within the limits of our simulation experiment, the collapsed and the underlying correlation are relatively close. Under the worst conditions of our simulation, where the underlying relation is very strong and the variables are collapsed into only two categories, the difference in correlations is about 0.2. Under all other conditions the difference is less. In some applications this inaccuracy may be intol-

[3] With skewed variables such as population and land area, highly influential observations may affect the correlations. One corrective response is to transform the variables to reduce skewness and to lessen the potential outlier effects. Because our interest is in the effect of collapsing on skewed variables, we do not transform the variables for our example.

erable, but for many exploratory areas of research such inaccuracies may be more readily accepted.

In contrast to O'Brien (1979), who found that the correlation of a collapsed variable with itself is a nonmonotonic function of the number of categories, we find the correlation between two different collapsed variables to be a *monotonic* function of the number of categories.[4] Since most researchers are interested in examining the interrelationship between distinct variables rather than a variable correlated with itself, our research supports the general view that "the more categories the better." In addition, our results suggest that at a minimum five or six categories should be used, since the collapsed variables' correlations become considerably closer to the correlation of the continuous variables when at least this number of scale points are available.

Several issues remain to be explored before generalizing our findings to empirical research. First, research is needed to determine which underlying distributions most often characterize the actual variables used in social research. Simulations and experiments with actual data that match these common distributions (if the continuous measures are available) would then be useful. Second, the possibility of correlated measurement error between variables collapsed in the same manner should be explored. Since much of our measurement theory is built on the assumption of independent errors of measurement, our usual corrections for attenuations and other results based on classical measurement theory may not be appropriate. Finally, the issue of the implications of collapsed variables for multivariate research deserves considerable attention.

REFERENCES

Acock, Alan C. and J. David Martin
1974 "The undermeasurement controversy: should ordinal data be treated as interval?" Sociology and Social Research 58:427–33.

Allen, Michael Patrick
1976 "Conventional and optimal interval scores for ordinal variables." Sociological Methods and Research 4:475–94. (See Errata in 1977, 5:436).

Baker, B. O., C. D. Hardyck, and L. F. Petrinovich
1966 "Weak measurement versus strong statistics." Educational and Psychological Measurement 26:291–309.

Banks, Arthur S. and R. B. Textor
1963 A Cross-Polity Survey. Cambridge: MIT Press.

Bendig, A. W.
1954 "Reliability and the number of rating scale categories." The Journal of Applied Psychology 38:38–40.

Benson, Purnell H.
1971 "How many scales and how many categories shall we use in consumer research? A comment." Journal of Marketing, October:59–61.

Blalock, Hubert M., Jr.
1979 "The presidential address: measurement and conceptualization problems: the major obstacle to integrating theory and research." American Sociological Review 44:881–94.

Champion, Dean J.
1968 "Some observations on measurement and statistics: a comment." Social Forces 46:541.

Green, Paul E. and V. R. Rao
1970 "Rating scales and information recovery—how many scales and response categories to use?" Journal of Marketing 34:33–9.
1971 "A rejoinder to 'How many scales and how many categories shall we use in consumer research?'—a comment." Journal of Marketing, October:61–2.

Grether, David M.
1976 "On the use of ordinal data in correlation analysis." American Sociological Review 41:908–12.

Henkel, R.
1975 "Part-whole correlations and the treatment of ordinal and quasi-integral data as interval data." Pacific Sociological Review 18:3–26.

Hensler, Carl and Brian Stipak
1979 "Estimating interval scale values for survey item response categories." American Journal of Political Science 23:627–49.

Jacoby, Jacob and Michael S. Matell
1971 "Three-point likert scales are good enough." Journal of Marketing Research 23:495–500.

Kim, Jae-On
1975 "Multivariate analysis of ordinal variables." American Journal of Sociology 81:261–98.
1978 "Multivariate analysis of ordinal variables revisited." American Journal of Sociology 84:448–56.

Komorita, S. S. and W. K. Graham
1965 "Number of scale points and the reliability of scales." Educational and Psychological Measurement 25:987–95.

[4] An exception to this monotonic relationship occurs in the second category case of our simulation when moving from five to six categories (see Table 1).

Labovitz, Sanford
1967 "Some observation on measurement and statistics." Social Forces 46:151–160.
1968 "Reply to Champion and Morris." Social Forces 46:543–4.
1970 "The assignment of numbers to rank order categories." American Sociological Review 35:515–24.
1971 "In defense of assigning numbers to ranks." American Sociological Review 36:521–22.
1972 "Statistical usage in sociology: sacred cows and rituals." Sociological Methods and Research 1:13–37.
1975 "Comment on Henkel's paper: the interplay between measurement and statistics." Pacific Sociological Review 18:27–35.

Lancaster, H. O. and M. A. Hamdan
1964 "Estimation of the correlation coefficient in contingency tables with possibly nonmetrical characters." Psychometrika 29:383–91.

Lehmann, Donald R. and James Hulbert
1972 "Are three-point likert scales always good enough?" Journal of Marketing Research 9:444–6.

Mayer, Lawrence S.
1970 "Comment on 'The assignment of numbers to rank order categories'." American Sociological Review 35:916–17.
1971 "A note on treating ordinal data as interval data." American Sociological Review 36:519–20.

Morris, Raymond N.
1968 "Some observations on measurement and statistics: further comment." Social Forces 46:541–2.

O'Brien, Robert M.
1979 "The use of Pearson's r with ordinal data." American Sociological Review 44:851–7.

Pearson, Karl
1904 "Mathematical contribution to the theory of evolution XIII: on the theory of contingency and its relation to association and normal correlation." Drapers Company Research Memoirs, Biometrics Series, Number 1.
1913 "On the measurement of the influence of 'broad' categories on correlation." Biometrika 9:116–39.

Pearson, Karl and David Heron
1913 "On theories of association." Biometrika 9:159–315.

Richie-Scott, A.
1918 "The correlation coefficient of a polychoric table." Biometrika 12:93–133.

Russett, Bruce M. et al.
1964 World Handbook of Political and Social Indicators. New Haven: Yale University.

Schweitzer, S. and D. G. Schweitzer
1971 "Comment on the Pearson r in random number and precise functional scale transformations." American Sociological Review 36:518–9.

Stern, Maxine S.
1971 "The effects of grouping on the correlation coefficient." Paper presented at the annual meeting of the American Sociological Association, Denver, Colorado.

Symonds, Percival M.
1924 "On the loss of reliability in ratings due to coarseness of the scale." Journal of Experimental Psychology 7:456–61.

Vargo, L. G.
1971 "Comment on 'The assignment of numbers to rank order categories'." American Sociological Review 36:517–8.

Wilson, T. P.
1971 "Critique of ordinal variables." Social Forces 49:432–44.

Wylie, Peter B.
1976 "Effects of coarse grouping and skewed marginal distributions on the Pearson product moment correlation coefficient." Educational and Psychological Measurement 36:1–7.