DOCUMENT RESUME

ED 052 263	24	TM 000 776

AUTHOR	Edwards, Keith J.
TITLE	Correcting Partial, Multiple and Canonical Correlations for the Attenuation of Measurement Error.
INSTITUTION	Johns Hopkins Univ., Baltimore, Md. Center for the Study of Social Organization of Schools.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	R-109
BUREAU NO	BR-6-1610
PUB DATE	Jul 71
GRANT	OEG-2-7-061610-0207
NOTE	23p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Algebra, Analysis of Variance, *Correlation, Data Analysis, Mathematical Models, Predictor Variables, *Reliability, *Standard Error of Measurement, *Statistical Analysis, Statistics, True Scores

ABSTRACT

This paper, a revision of the original document, "Correcting Partial, Multiple, and Canonical Correlations for Attenuation" (see TM 000 535), presents the formula for correcting coefficients of partial correlation for attenuation due to errors of measurement. In addition, the correction for attenuation formulas for multiple and cannonical correlations are discussed. The attenuating effects of measurement error are examined algebraically for a special case of partial correlation and multiple correlation. Finally, the formula for the corrected partial r is related to recent work on the measurement of change. (Author/AG)

BR 6-1610
PA 24
TM

Center for Social Organization of Schools

ED052263

REPORT No. 109

JULY, 1971

CORRECTING PARTIAL, MULTIPLE AND CANONICAL CORRELATIONS FOR THE ATTENUATION OF MEASUREMENT ERROR

KEITH J. EDWARDS

The Johns Hopkins University

TM 000 729

STAFF

John L. Holland, Director

James M. McPartland, Assistant Director

Virginia Bailey
Thelma Baldwin
Zahava D. Blum
Judith P. Clark
James S. Coleman
David DeVries
Keith Edwards
Doris R. Entwisle
Gail Fennessey
James Fennessey
Jean Finnegan
Ellen Greenberger
Rubie Harris
Edward J. Harsch
Robert T. Hogan
John H. Hollifield
Nancy L. Karweit
Judith Kennedy
Steven Kidder
Hao-Mei Kuo
Samuel Livingston
Edward L. McDill
Rebecca J. Muraro
Jeanne O'Connor
Suzanne K. Pieper
Peter H. Rossi
Leslie Schnuelle
Aage B. Sørensen
Julian C. Stanley
Mary C. Viernstein
Diana F. Ward
Murray A. Webster
Phyllis K. Wilson

CORRECTING PARTIAL, MULTIPLE AND CANONICAL CORRELATIONS FOR THE ATTENUATION OF MEASUREMENT ERROR

GRANT NO. OEG-2-7-061610-0207

PROJECT NO. 61610-01

KEITH J. EDWARDS

REPORT NO. 109

JULY, 1971

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred.

The Johns Hopkins University
Baltimore, Maryland

INTRODUCTORY STATEMENT

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through five programs to achieve its objectives. The Academic Games program has developed simulation games for use in the classroom, and is studying the processes through which games teach and evaluating the effects of games on student learning. The Social Accounts program is examining how a student's education affects his actual occupational attainment, and how education results in different vocational outcomes for blacks and whites. The Talents and Competencies program is studying the effects of educational experience on a wide range of human talents, competencies, and personal dispositions in order to formulate-- and research--important educational goals other than traditional academic achievement. The School Organization program is currently concerned with the effects of student participation in social and educational decision-making, the structure of competition and cooperation, formal reward systems, ability-grouping in schools, and effects of school quality. The Careers and Curricula program bases its work upon a theory of career development. It has developed a self-administered vocational guidance device to promote vocational development and to foster satisfying curricular decisions for high school, college, and adult populations.

This report, like others occasionally published by the Center, deals with a subject common to all programs -- that of scientific measurement.

Acknowledgment

This paper is based on a paper presented at the annual meeting of the American Educational Research Association, New York, New York, February 5, 1971. I thank Gene V. Glass and Julian C. Stanley for their helpful comments on the original paper.

ABSTRACT

The formula for correcting coefficients of partial correlation for attenuation due to errors of measurement is derived. In addition, the correction for attenuation formulas for multiple and canonical correlations are presented and discussed. The attenuating effects of measurement error are examined algebraically for a special case of partial correlation and by means of numerical examples for the general case of partial correlation and multiple correlation. Finally, the formula for the corrected partial r is related to recent work on the measurement of change.

Introduction

In dealing with multivariate correlational techniques and fallible data, one is faced with the same difficulties that have been pointed out for the product-moment correlation (Finucci, 1971) and other similar measures of association (Stanley, 1971). Correlations based on fallibly measured variables will result in values which are underestimates of the correlations among the true parts of the variables. (The truth of this statement for all multivariate situations has not been proven analytically, but Cochran's (1970) work regarding multiple correlation suggests that the statement does, in fact, hold true.) Investigators have been inclined to ignore the problems of unreliability, being content with underestimates of the true relationship and avoiding "questionable" correction for attenuation procedures. However, such an approach ignores useful information. In addition to providing a means for obtaining estimates of true-score correlations, correction for attenuation formulas facilitate understanding of the effects of unreliability on the results. Information of this type is useful, for example, in deciding how much could be gained by expending time and money to develop more reliable measurement.

The purpose of the present paper is to give the correction for attenuation formulae for partial, multiple, and canonical correlation coefficients and to discuss, where known, the effects of measurement error on these statistics. Most of the formulas presented have been derived elsewhere in the literature. I have simply standardized the notation and extended some of the derivation where appropriate.

Partial Correlation Corrected for Attenuation

First, consider the first-order partial correlation coefficient. Suppose there are three fallibly measured variables x_1, x_2, and x_3 of, say, alienation, school achievement, and I.Q. and one wants to know the true correlation between alienation and school achievement, controlling for I.Q. Begin by defining the variable x_i to be the sum of its true score, t_i, and errors of measurement, e_i ($x_i = t_i + e_i$ for i = 1, 2, 3). (In the sequel it is assumed that $E(x_i) = 0$.) We assume that the true score and error for each variable covary zero ($\sigma_{t_i e_i} = 0$); that the true score of one variable and the error component in another variable covary zero ($\sigma_{t_i e_j} = 0$); and that the errors in the differential variables covary zero ($\sigma_{e_i e_j} = 0$).

The partial correlation between alienation and achievement, controlling for I.Q., is defined as the zero-order correlation of residuals. The residuals for alienation are given by the difference between the observed values and the regression estimates of alienation from I.Q. The residuals are represented symbolically in equations (1) and (2). It is well known that

$$x_{1.3} = x_1 - \beta_{x_1 x_3} x_3 \tag{1}$$

$$x_{2.3} = x_2 - \beta_{x_2 x_3} x_3 \tag{2}$$

the correlation of residuals can be expressed in terms of the three zero-order correlations. The formula is given in (3), which is the partial r

$$\rho_{x_1x_2\cdot x_3} = \frac{\sigma_{x_{1.3}x_{2.3}}}{\sqrt{\sigma^2_{x_{1.3}} \cdot \sigma^2_{x_{2.3}}}} = \frac{\rho_{x_1x_2} - \rho_{x_1x_3}\rho_{x_2x_3}}{\sqrt{(1 - \rho^2_{x_1x_3})(1 - \rho^2_{x_2x_3})}} \tag{3}$$

based on fallible measures. The partial correlation coefficient, corrected for attenuation, would yield the partial correlation of true score, i.e., the correlation of true score residuals. We can obtain the correction for attenuation formula by starting with the correlation of true score residuals and working backwards.

The true score residuals are defined as the difference between the true value and the estimated true value based on a regression of the variable (t_1 or t_2) on the true value of the control variable (t_3) and are given in formulas (4) and (5). The partial correlation of t_1 and t_2 controlling for t_3

$$t_{1.3} = t_1 - \beta_{t_1t_3}t_3 \tag{4}$$

$$t_{2.3} = t_2 - \beta_{t_2t_3}t_3 \tag{5}$$

is then given by (6). Expanding numerator and denominator, we can use some

$$\rho_{t_1t_2\cdot t_3} = \frac{\sigma_{t_{1.3}t_{2.3}}}{\sqrt{\sigma^2_{t_{1.3}} \cdot \sigma^2_{t_{2.3}}}} = \frac{\sigma_{(t_1 - \beta_{t_1t_3}t_3)(t_2 - \beta_{t_2t_3}t_3)}}{\sqrt{\sigma^2_{(t_1 - \beta_{t_1t_3}t_3)}\sigma^2_{(t_2 - \beta_{t\,t_3}t_3)}}} \tag{6}$$

of the well-known properties of classical test theory to express the true partial correlation in terms of the attenuated zero-order correlations and reliabilities. The result, given in (7) is the correction for attenuation

$$\rho_{t_1t_2.t_3} = \frac{\rho_{33}\rho_{x_1x_2} - \rho_{x_1x_3}\rho_{x_2x_3}}{\sqrt{\left(\rho_{11}\rho_{33} - \rho^2_{x_1x_2}\right)}\sqrt{\left(\rho_{22}\rho_{33} - \rho^2_{x_2x_3}\right)}} \tag{7}$$

formula for a first-order partial correlation. (The proof is given in the appendix.) It should be noted that the formula is equivalent to correcting each of the zero-order correlations for attenuation by the usual way and substituting these values into (3). (See Livingston and Stanley, 1970.)

A General Approach to Multivariate Corrections for Attenuation

Meredith (1964) developed a more general approach to correction for attenuation problems which he has applied to the canonical correlation problem. His result can be readily applied to problems involving partial and multiple correlation. We begin with a variance-covariance matrix, Σ_x, of rank p + q, where p + q is the number of variables. Under the assumption that the classical test theory model is appropriate for each of the p + q variables we can write the matrix Σ_x as the sum of two matrices, Σ_t, the variance-covariance matrix among true scores, and Σ_e, the variance-covariance matrix among the errors of measurement (equation 8).

$$\Sigma_X = \Sigma_T + \Sigma_E \tag{8}$$

Assuming errors of measurement covary zero with each other the matrix Σ_e is a diagonal matrix of the variance errors of estimates. We can obtain Σ_t by subtraction (equation 9). Given Σ_t, the matrix of true

$$\Sigma_T = \Sigma_X - \Sigma_E \tag{9}$$

score variances and covariances, it is a simple matter to obtain the matrix of true score correlations by dividing each element by the square root of the product of the appropriate variances. These operations are shown in matrix notation in equation (10).

$$R_T = D(\Sigma_T)^{-\frac{1}{2}} \Sigma_T D(\Sigma_T)^{-\frac{1}{2}} \tag{10}$$

It is important to note here that (10) is equivalent to correcting each of the zero-order correlations in R_x, the matrix of observed correlations, for attenuation in the usual manner. That such is the case becomes clear if we consider each of the p + q variables to have a mean of 0 and a variance of 1. Under the latter condition $\Sigma_x = R_x$ and Σ_e is a diagonal matrix of alienation coefficients. Thus, the matrix Σ_t of (9) is the matrix of observed intercorrelations (R_x) with reliabilities on the diagonal, which is the true-score variance-covariance matrix of standardized variables. The operations shown in (10) now involve dividing every correlation in R_x by the square root of the product of the reliabilities for the appropriate variables, which is the zero-order correction for attenuation procedure.

So far, the discussion has been in terms of population values. Meredith has pointed out that a maximum likelihood estimate of Σ_t and thus of R_t can be obtained from $\hat{\Sigma}_x$, the sample variance-covariance matrix, if the reliabilities of the measures are known (equations 11 and 12). Though the remainder of the paper continues to use the population values, one can easily substitute $\hat{R}_t$ under the above restriction.

$$\hat{\Sigma}_T = \hat{\Sigma}_X - \Sigma_E \qquad (11)$$

$$\hat{R}_T = D(\hat{\Sigma}_T)^{-1/2}\hat{\Sigma}_T D(\hat{\Sigma}_T)^{-1/2} \qquad (12)$$

A general procedure for correcting multivariate correlations for attenuation involves the following two steps. <u>First</u>, correct each of the zero-order correlations for attenuation in the usual way to obtain R_t. <u>Second</u>, calculate the desired statistic from R_t.

Let us return to the problem of partial correlations. Suppose that we were interested in obtaining the true score correlations among a set of p variates controlling for true scores on a second set of q variables. We could solve the problem by first obtaining R_t, partitioning R_t as shown in (13), and using the matrix solution for partial correlations (Anderson,

$$R_T = \left[\begin{array}{c|c} R_{T_{11}} & R_{T_{12}} \\ \hline R_{T_{21}} & R_{T_{22}} \end{array}\right] \tag{13}$$

$$R_{T_{1.2}} = R_{T_{11}} - R_{T_{12}} R_{T_{22}}^{-1} R_{T_{21}} \tag{14}$$

1958:29 and Morrison, 1967:89) shown in (14) above. If q = 1, $R_{T_{1.2}}$ is a p x p matrix of first order partials whose off-diagonal elements are of the form given in (3).

The multiple correlation problem involves finding the maximum correlation between a single criterion and a linear combination of, say, p predictors. The matrix solution for squared multiple correlation (MR^2) is given in (15) (See Anderson, 1958: 30, or Morrison, 1967:104). The MR^2 between the true scores of the p predictors and the criterion could be

$$MR_X^2 = R_{X_{21}} R_{X_{11}}^{-1} R_{X_{12}} \tag{15}$$

obtained by substituting the corresponding true score correlation matrices of (13) into (15), resulting in equation (16). ($R_{T_{21}}$ and $R_{T_{12}}$ are vectors.)

$$MR_T^2 = R_{T_{21}} R_{T_{11}}^{-1} R_{T_{12}} \tag{16}$$

In the above situation, any of the p + 1 variables could be designated as the criterion by simply interchanging the appropriate rows and columns of R_t. A general formula for $MR^2_{T_i}$ of each of the i variates with the remaining variates corrected for attenuation is given by (17), where I is a p + 1 identity matrix and D indicates diagonals of the matrices given in parentheses.

$$D(MR^2_{T_i}) = I - D^{-1}(P_T^{-1}) \qquad (17)$$

The last statistic we shall discuss is the canonical correlation coefficient (Hotelling, 1936). Canonical correlation is a generalization of the concept of multiple correlation to the case of multiple criteria (q > 1) as well as multiple predictors (p > 1). The objective in such an analysis is to find the maximum correlation between a linear composite of the predictors and a linear composite of the criteria. Though Hotelling was primarily concerned with the largest correlation between these composites, there are k = min (p, q) possible independent correlations. The k canonical correlations for any given set of p predictors and q criteria are given by the latent roots of $R_{21}R_{11}^{-1}R_{12}R_{22}^{-1}$. Meredith (1964) has shown that if the true-score correlation matrix in (13) is used, you would have the k canonical correlations corrected for attenuation. The weighting vectors to form the linear composite of the criterion variates and the linear composite of predictor variables are the latent vectors of $R_{21}R_{11}^{-1}R_{12}R_{22}^{-1}$ and $R_{12}R_{22}^{-1}R_{21}R_{11}^{-1}$. The formulas would provide either the attenuated weighting vectors or the true-score weights, depending on which correlation matrices were used.

Effects of Errors of Measurement

In the introduction it was pointed out that the most well-known effect of errors of measurement on zero-order correlations is to produce an underestimate of the true value. This appears to also be the case for multiple correlation. For example, Cochran (1970) showed that for a number of situations involving multiple correlation, a good estimate of attenuating effects of fallible data is given by (18). Though the actual value of MR_X^2 may run 25% higher than this value for positive ρ_{iq} and low predictor reliability ($\rho_{ii} = 0.5$), it never exceeds MR_T^2.

$$MR_X^2 = MR_T^2 \cdot \rho_{qq} \frac{\sum_{i=1}^{q-1} \rho_{iq}\rho_{ii}}{\sum_{i=1}^{q-1} \rho_{iq}} = R_T^2 \rho_{qq} \bar{\rho}_{ii} \qquad (18)$$

where ρ_{qq} is reliability of the criterion

ρ_{ii} is reliability of the ith predictor

ρ_{iq} is correlation between criterion and ith predictor.

Bohrnstedt (1969) derived a formula for correcting partial correlations for attenuation due to errors of measurement which is similar to (7), but does not contain the terms ρ_{11}, ρ_{22}. Upon examining his derivation, it was apparent that he was correcting only for errors of measurement in the control variable x_3. In effect, he had provided the formula for a partially corrected partial correlation coefficient. On the basis of his formula, Borhnstedt indicates that it is possible for the corrected partial correlation to be less than the obtained partial correlation. While this appears to be the case the author's verbal argument is misleading. He points out that in his formula, given in (19), the factor ρ_{33} in the numerator would

$$\rho_{x_1x_2.t_3} = \frac{\rho_{33}\rho_{x_1x_2} - \rho_{x_1x_3}\rho_{x_2x_3}}{\sqrt{(\rho_{33}-\rho_{x_1x_3}^2)(\rho_{33}-\rho_{x_2x_3}^2)}} \tag{19}$$

tend to make (19) smaller than (3). (Bohrnstedt, 1969:127). He does not mention that the presence of ρ_{33} in the deonominator of (19) would tend to make (19) larger than (3). The net effect of these two factors is not obvious. By making the simplifying assumption that $\rho_{13} = \rho_{23}$ (which is not very restrictive), we can show algebraically that (19) will always be less than (3). Assuming the correlations with the control variable to be equal, (3) and (19) reduce to

$$\rho_{x_1x_2.x_3} = \frac{\rho_{x_1x_2} - \rho_{x_1x_3}^2}{(1 - \rho_{x_1x_3}^2)} \tag{3'}$$

$$\rho_{x_1x_2.t_3} = \frac{\rho_{x_1x_2} - \rho_{x_1x_3}^2/\rho_{33}}{(1 - \rho_{x_1x_3}^2/\rho_{33})} \tag{19'}$$

Letting the quantities in (3') be represented by $\frac{a - b^2}{(1 - b^2)}$ and the quantities in (19') by $\frac{a - d^2}{(1 - d^2)}$, it is sufficient to show that

$$G = \frac{a - b^2}{1 - b^2} - \frac{a - d^2}{1 - d^2} > 0$$

Creating a common denominator for G one secures:

$$G = \frac{(1 - d^2)(a - b^2) - (1 - b^2)(a - d^2)}{(1 - b^2)\ (1 - d^2)}$$

Since the denominator will always be positive, it is sufficient to show that

$$G^* = (a - ad^2 - b^2 + d^2b^2) - (a - ab^2 - d^2 + b^2d^2) > 0$$

Simplifying the above expression yeilds

$$G^* = (d^2 - b^2) - a(d^2 - b^2).$$

The value d represents the zero order correlation between variables 1 and 3 corrected for errors of measurement in x_3 and b represents the attenuated correlation. It is well known that $d > b$. Thus, since $|a| < 1$, it is true that $G^* > 0$ which completes the proof that $\rho_{x_1x_2.x_3} > \rho_{x_1x_2.t_3}$. The proof of this latter inequality when $\rho_{x_1x_3} \neq \rho_{x_2x_3}$ becomes algebraically intractable. However, one effect of correcting for attenuation due to a fallible x_3 only, is to decrease the size of the residuals being correlated in (19). It is well known that a restriction in the range of a variable decreases correlations involving the variable. As Bohrnstedt points out, "the smaller the residual variation to correlate, the smaller the partial correlation (127)."

Blalock (1964: 149-150) has noted that it is possible to observe a non-zero partial correlation when the true relationship is zero, due to errors of measurement in the control variable. That such is the case is clear only when $\rho_{11} = \rho_{22} = 1$ and thus (19) is the true partial correlation. However, when all three variates are fallibly measured, the size of the observed partial relative to the true partial is difficult to determine algebraically. The information presented in Table 1 does provide some information on their relative values for a range of reliabilities and correlations. In order to simplify computation and presentation it was assumed that $\rho_{11} = \rho_{22}$ and $\rho_{x_1x_3} = \rho_{x_2x_3}$. Out of the ten true partial correlations possible in Table 1, three of them were <u>less</u> than the attenuated partial r. One of

TABLE 1

Comparison of $\rho_{t_1t_2 \cdot t_3}$ and $\rho_{x_1x_2 \cdot x_3}$ for

Different Values of $\rho_{11}, \rho_{22}, \rho_{33}, \rho_{x_1x_2}, \rho_{x_1x_3}$, And $\rho_{x_2x_3}$

RELIABILITIES ($\rho_{33}, \rho_{22}, \rho_{11}$)	CORRELATIONS ($\rho_{13}, \rho_{23}, \rho_{12}$) (.2, .2, .3)	(.5, .5, .3)	(.2, .2, .7)	(.5, .5, .7)	(.6, .6, .7)
(.3, .4, .4)	.63*	**	**	**	**
(.9, .4, .4)	.72	.18	**	**	**
(.3, .8, .8)	.25†	**	.86	**	**
(.9, .8, .8)	.34	.04†	.87	.81	.50†
$\rho_{x_1x_2 \cdot x_3}$ =	.27	.07	.69	.60	.53

* Numbers in the Table are $\rho_{t_1t_2 \cdot t_3}$

** Impossible because $\rho_{t_1t_2 \cdot t_3} > 1$

† True value is less than the attenuated value.

these occurred in the cell where the reliability of the control variable was small (.30) and the reliability of the other variables was large (.80). The other two cases occurred in cells where all reliabilities were high (.90, .80) and the correlations with the control variable were large (.60). The data in Table 1 suggest that when all three variables are fallibly measured, the attenuated partial r will more often be an underestimate of the true value than an overestimate and that the underestimation is more severe than the overestimation. These tentative conclusions of course need to be verified by more extensive monte carlo simulations over a wider range of values.

It may be useful at this point to relate equation (7) to recent work on the measurement of change by Tucker, Damarin, and Messick (1966) and Cronback and Furby (1970)(hereafter referred to as TDM and CF, respectively). Of primary interest in these papers is the definition of a "true" residual change score. The two equations, in a form consistent with this paper, are

$$\hat{t}_{1.3} = x_1 - \beta_{t_1 t_3} x_3 \qquad (20)\ \text{(TDM's Equation 25)}$$

$$t_{1.3} = t_1 - \beta_{t_1 t_3} t_3 \qquad (21)\ \text{(CF's Equation 22)}$$

(Where x_3 and x_1 are the pretest and posttest measures, respectively.) The formula for residual in (20) by TDM is similar to Bohrnstedt's work in that it only takes into account the reliability of the pretest (control) variable. The equation in (20) used by CF is identical to (4). Both papers propose the use of the residual change score for correlation with other variables (TDM, 1966: 470 and CF, 1970:76). The resulting part correlations of either $\hat{t}_{1.3}$

or $t_{1.3}$ with x_2 are attenuated by the fallible x_2, since ρ_{22} is not taken into account.

A second problem raised when errors of measurement are present is the valid interpretation of results for multivariate correlations. In multiple and canonical correlation studies, an important objective is to discover the relative importance of the predictor and criterion variables. The intercorrelations among these variables and their unreliability can interact to produce misleading results. An example from Cochran (1970:33) illustrates this point.

A common practice in the application of multiple correlation (especially among sociologists) is to partition the predicted variance MR^2 into portions uniquely attributable to each predictor and the portion of common variance predicted. Unreliability can have a substantial effect on the results of such an analysis. Consider the 2-predictor case with ρ_{13} = ρ_{23} = 0.5, and ρ_{12} = 0.3 and no error of measurement:

$MR^2_{3.12} = .385$

% of variance unique to x_1 = 13.5
% of variance unique to x_2 = 13.5
% of variance common to both = 11.5

(Uniqueness of x_1 is here defined as the drop in MR^2 that would result from deleting x_1 from the equation. Darlington (1968: 162) defines this as the "usefulness" of a variable. It is clear in the above example that $MR^2_{3.2} = \rho^2_{23} = .25$, $MR^2_{3.12} - MR^2_{3.2} = .135$, the uniqueness for x_1.)

With the reliability of variable 1 equal to .8; i.e. $\rho_{11} = 0.8$ and $\rho_{22} = 1$,

$MR^2_{3.12} = .356$

% of variance unique to x_1 = 10.6
% of variance unique to x_2 = 15.6
% of variance common to both = 9.4

With $\rho_{11} = 0.6$ and $\rho_{22} = 1$

$MR^2_{3.12} = .328$

% of variance unique to x_1 = 7.8
% of variance unique to x_2 = 17.8
% of variance common to both = 7.2

In the above example we see that as the reliabilities of the predictors become more disparate, the true contribution of each variable becomes more distorted. This effect can be best understood when one considers what would happen if x_1 were removed from the correlation, altogether. $MR^2_{3.2}$ would equal .25 and the percent of variance unique to x_2 would be 25. Unreliability in one of the variables takes part of that variable "out" of the prediction, shifting predicted variance to the more reliable predictors. The change in the importance of predictors in multiple correlation caused by deletion of one of the variables has been referred to as the "bouncing betas." It is apparent from these examples that differences in the reliabilities adds more bounce to these results. These results on the effects of disparate reliabilities are one more reason to heed the warning given by Darlington, "It would be better to simply concede that the notion of 'independent contribution to variance' has no meaning when predictor variables are intercorrelated." (1968:169)

References

Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons Inc., 1958.

Blalock, H. M. *Causal Inference in Nonexperimental Research*. Chapel Hill: University of North Carolina Press, 1964.

Bohrnstedt, G. W. "Observations on the Measurement of Change" in *Sociological Methodology*, Edgar F. Borgatta (ed.) San Francisco: Jossey-Bass, Inc., 1969. pp. 113-133.

Cochran, W. G. "Some Effects of Errors of Measurement on Multiple Correlation" in *Journal of the American Statistical Association*, 1970, 65, 22-34.

Cronbach, L. J. and Furby, L. "How We Should Measure "Change"--Or Should We." *Psychological Bulletin*, 1970, 74, 68-80.

Darlington, Richard B. "Multiple Regression in Psychological Research and Practice." *Psychological Bulletin*, 1968, 69, 161-182.

Finucci, J. M. "Early History and Basic Derivation of Attenuation Formulas." A paper read at the American Educational Research Association Convention. New York: February 5, 1971.

Hotelling, H. "Relations Between Two Sets of Variates," *Biometrika*, 1936, 28, 321-377.

Livingston, S. A. and Stanley, J.C. "What is the True Coefficient of Partial Correlation?" Baltimore: Johns Hopkins University, Center for Social Organization of Schools. Report No. 107, July, 1971.

Meredith, W. "Canonical Correlations With Fallible Data, " *Psychometrika* 1964, 29, 55-65.

Morrison, D. *Multivariate Statistical Methods*. New York: McGraw-Hill, 1967.

Stanley, J. C. "Correcting Four Similar Correlational Measures for Attenuation due to Errors of Measurement in the Dependent Variable: η^2, ϵ^2, ω^2, and Intraclass r." *Educational and Psychological Measurement*, (in press).

Tucker, L. R. Damarin, F., and Messick, S. "A Base-Free Measure of Change." *Psychometrika*, 1966, 31, 457-473.

APPENDIX

Derivation of correction formula for partial correlation coefficient.

$$\rho_{t_1t_2\cdot t_3} = \frac{\sigma_{t_{1.3}t_{2.3}}}{\sqrt{\sigma^2_{t_{1.3}} \cdot \sigma^2_{t_{2.3}}}} = \frac{\sigma_{(t_1 - \beta_{t_1t_3}t_3)(t_2 - \beta_{t_2t_3}t_3)}}{\sqrt{\sigma^2_{(t_1 - \beta_{t_1t_3}t_3)}\sigma^2_{(t_2 - \beta_{t_2t_3}t_3)}}} \tag{1}$$

Expanding the numerator of (1)

$$\sigma_{t_{1.3}t_{2.3}} = \sigma_{t_1t_2} - \beta_{t_1t_3}\sigma_{t_2t_3} - \beta_{t_2t_3}\sigma_{t_1t_3} + \beta_{t_1t_3}\beta_{t_2t_3}\sigma^2_{t_3}$$

$$= \sigma_{t_1}\sigma_{t_2}\rho_{t_1t_2} - \left(\rho_{t_1t_3}\frac{\sigma_{t_1}}{\sigma_{t_3}}\right)\sigma_{t_2}\sigma_{t_3}\rho_{t_2t_3} - \left(\rho_{t_2t_3}\frac{\sigma_{t_2}}{\sigma_{t_3}}\right)\sigma_{t_1}\sigma_{t_3}\rho_{t_1t_3}$$

$$+ \left(\rho_{t_1t_3}\frac{\sigma_{t_1}}{\sigma_{t_3}}\right)\left(\rho_{t_2t_3}\frac{\sigma_{t_2}}{\sigma_{t_3}}\sigma^2_{t_3}\right)$$

(The last two terms are equivalent except for sign, and thus they sum to zero.)

$$= \sigma_{t_1}\sigma_{t_2}\rho_{t_1t_2} - \rho_{t_1t_3}\sigma_{t_1}\sigma_{t_2}\rho_{t_2t_3}$$

$$= \sqrt{\rho_{11}\rho_{22}}\sigma_{x_1}\sigma_{x_2}\frac{\rho_{x_1x_2}}{\sqrt{\rho_{11}\rho_{22}}} - \frac{\rho_{x_1x_3}}{\sqrt{\rho_{11}\rho_{33}}}\left(\sqrt{\rho_{11}\rho_{22}}\sigma_{x_1}\sigma_{x_2}\right)\frac{\rho_{x_2x_3}}{\sqrt{\rho_{22}\rho_{33}}} \quad ,$$

Therefore, where ρ_{11} is the reliability coefficient of variable x_1

$$\sigma_{t_{1.3}t_{2.3}} = \sigma_{x_1}\sigma_{x_2}\left(\frac{\rho_{33}\rho_{x_1x_2} - \rho_{x_1x_3}\rho_{x_2x_3}}{\rho_{33}}\right) \quad . \tag{2}$$

Expanding the first term in the denominator of (1) one secures

$$\sigma^2_{t_{1.3}} = \sigma^2_{t_1} + \beta^2_{t_1t_3}\sigma^2_{t_3} - 2\beta_{t_1t_3}\sigma_{t_1t_3}$$

$$= \sigma^2_{t_1} + \rho^2_{t_1t_3}\sigma^2_{t_1} - 2\rho_{t_1t_3}\sigma^2_{t_1}\rho_{t_1t_3}$$

$$= \sigma^2_{t_1} - \rho^2_{t_1t_3}\sigma^2_{t_1}$$

$$= \rho_{11}\sigma^2_{x_1} - \frac{\rho^2_{x_1x_3}\rho_{11}\sigma^2_{x_1}}{\rho_{11}\rho_{33}} \quad .$$

Therefore,

$$\sigma^2_{t_{1.3}} = \sigma^2_{x_1} \frac{\left(\rho_{11}\rho_{33} - \rho^2_{x_1x_3}\right)}{\rho_{33}} \quad . \tag{3}$$

Similarly, it can be shown that

$$\sigma^2_{t_{2.3}} = \sigma^2_{x_2} \frac{\left(\rho_{22}\rho_{33} - \rho^2_{x_2x_3}\right)}{\rho_{33}} \quad . \tag{4}$$

Substituting (2), (3), and (4) into (1) and simplifying we get

$$\rho_{t_1t_2 \cdot t_3} = \frac{\rho_{33}\rho_{x_1x_2} - \rho_{x_1x_3}\rho_{x_2x_3}}{\sqrt{\left(\rho_{11}\rho_{33} - \rho^2_{x_1x_3}\right)}\sqrt{\left(\rho_{22}\rho_{33} - \rho^2_{x_2x_3}\right)}} \quad .$$

which is the formula for the partial correlation coefficient corrected for attenuation.