An Empirical Study of the Accuracy of Corrections for Restriction in Range Due to Explicit Selection

Jack M. Greener and H. G. Osburn
University of Houston

An empirical study of the corrections for restriction in range due to explicit selection resulted in the following conclusions. (1) The corrected sample correlation was no more accurate than the uncorrected sample correlation for low unrestricted population correlations in the range .10 to .25. (2) For large unrestricted population correlations in the range .60 to .80, the corrected sample correlation was always more accurate than the uncorrected sample correlation. (3) For moderate (.30 to .55) unrestricted population correlations, the corrected sample correlation was typically more accurate than the uncorrected sample correlation. (4) The correction was very sensitive to moderate departures from linearity but was quite robust in the face of rather substantial departures from homoscedasticity.

A serious methodological problem in test validation studies is restriction in range on the predictor due to explicit selection. Pearson (1903) offered an apparent solution to the restriction in range problem by deriving correction formulas based on a limited number of assumptions: (1) linearity of regression of Y on X; (2) homoscedasticity of error distributions; and (3) normally distributed variables. It was later discovered that the normality assumption is not required to develop the correction formula (Lawley, 1943–4).

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 3, No. 1 Winter 1979 pp. 31-41

Rydberg (1963) reviewed several studies which indicated that (1) corrected correlations were substantially better estimates of population values compared with restricted correlations; (2) there was no consistent tendency to overcorrect or undercorrect; and (3) the corrected correlations were approximately distributed within the bounds of normal sampling error. Additional research has shown that the decision as to which predictor or predictors would be best in a particular situation is often different when comparing uncorrected values, as opposed to corrected values, and that the decision based on the corrected values produces better results (Meredith, 1958; Srinivasan & Weinstein, 1973). In contrast to Rydberg's conclusions, Forsyth (1971) found that the traditional standard error for the Pearson product-moment coefficient is not adequate for establishing confidence intervals for corrected correlation coefficients.

Although past research has suggested that corrected correlations tend to be more accurate than uncorrected sample correlations, a number of investigators have questioned the accuracy of the correction under certain conditions. Lord and Novick (1968) pointed out that there is a tendency for test score data to violate both linearity and homoscedasticity assumptions at both extremes of the distribution and expressed serious reservations regarding the probable accuracy of the corrections under conditions of extreme selection. Also, in a large-scale empirical study investigating the effects of moderate to severe selection for a variety of test score data, Novick and Thayer (1969) concluded that the correction formulas were generally unsatisfactory, even for situations involving moderate selection. They were especially concerned with a tendency of the formulas to overcorrect for some types of data.

In a recent monte carlo study of the accuracy of the correction formula for explicit selection, it was found that corrected correlations were more accurate than uncorrected correlations for small degrees of truncation. This was true for a variety of simulated distributions that systematically violated the linearity and homoscedasticity assumptions. However, serious problems of overestimation occurred in some distributions with high degrees of truncation (Greener & Osburn, 1977).

The purpose of the present study was to examine the accuracy of the correction formula for explicit selection in a wide variety of empirical distributions that contained violations of either the linearity or homoscedasticity assumption or both. Data were obtained on a wide variety of psychological measures in order to enhance the generalizability of the results as much as possible.

Method

Data

The empirical data were scores on a variety of psychological measures including cognitive abilities, biographical information, temperament, interests, academic achievement, and job performance for three different samples. These samples included (1) male and female high school students in Grades 9 through 12 from Project TALENT ($N = 13{,}478$); (2) applicants for hourly maintenance and operating positions in a petroleum refinery ($N = 2{,}250$); and (3) white collar managerial and professional employees of a large international petroleum company ($N = 5{,}900$).

TALENT variables examined include Information I (*R*–190) a measure of general information distributed across 15 areas, representing a broad spectrum of subjects available in many high schools; English Total (*R*–230), a measure of overall achievement in English usage; Reading Comprehension (*R*–250), a measure of ability to comprehend written materials; Creativity (*R*–260), a measure of ability to find ingenious solutions to practical problems; Abstract Reasoning (*R*–290), a nonverbal test measuring ability to find logical relationships among elements; Mathematics Total (*R*–340), a comprehensive test of high school mathematics achievement; Mature Personality (*R*–610), a scale from a self-report personality inventory; and Physical Science, Engineering, and Mathematics Interest (*P*–704), a measure of scientific interest. Exclusion of cases with missing data reduced the TALENT sample to approximately 10,500 for most analyses. Additional information regarding Project TALENT is available from American Institutes for Research (1972).

The refinery data provided for an examination of the relationships among the following aptitude tests: Test of Learning Ability (Richardson, Bellows, Henry, & Co.); Test of Chemical Comprehension (Richardson, Bellows, Henry, & Co.); and Mechanical Comprehension Test (Psychological Corporation). Variables examined in the white collar sample included Miller Analogies Test (Psychological Corporation); Non-Verbal Reasoning Ability Test (Richardson, Bellows, Henry, & Co.); Guilford-Zimmerman Temperament Survey (Sheridan Supply Co.); a biographical information inventory developed by the company; and a performance rating appraising the employee's overall effectiveness. Both the Guilford-Zimmerman Temperament Survey scale (GZTS Success) and the Biographical Inventory were scored with empirical keys to predict managerial success.

Whenever possible, criterion variables were selected to resemble those typically used in practice, such as achievement tests and preformance appraisals. Table 1 provides an overview of

Table 1

Data on Bivariate Distributions

Distribution	Array Standard Deviations			Linearity of Regression	
	Sig.*	Range	Trend	Sig.*	Trend
Refinery Sample (N = 2250)					
Chemical Comp. vs. Learning Ability	.01	4.5 to 6.1	No Trend	.01	Flat in lower tail
Mechanical Comp. vs. Learning Ability	.01	7.3 to 9.1	No Trend	NS	Flat in both tails
Mechanical Comp. vs. Chemical Comp.	.01	7.4 to 10.1	No Trend	.01	Flat in lower tail
White Collar Sample (N - 5900)					
Appraisal vs. Miller Analogies	NS	6.0 to 8.5	No Trend	NS	No Trend
Appraisal vs. Nonverbal Reasoning	NS	6.5 to 8.5	No Trend	NS	No Trend
Appraisal vs. GZTS Success	NS	6.0 to 8.0	No Trend	NS	No Trend
Appraisal vs. Biographical Inventory	NS	5.5 to 9.1	No Trend	NS	Flat lower tail, steep upper tail
TALENT Sample (N = 10,500)					
Information I vs. Reading Comp.	.01	19.3 to 25.0	No Trend	.01	Flat lower tail, steep upper tail
Information I vs. Mature Personality	.01	39.0 to 32.0	Decreasing	.05	Flat lower tail, steep upper tail
English Total vs. Reading Comp.	.01	15.0 to 7.0	Decreasing	.01	Slightly steeper in both tails
English Total vs. Creativity	.01	15.0 to 8.0	Decreasing	.01	Steep lower tail, flat upper tail
Math Total vs. Abstract Reasoning	.01	3.7 to 9.9	Increasing	.01	Flat lower tail, steep upper tail
Math Total vs. Scientific Interest	.01	6.0 to 11.0	Increasing	.01	Steeper in upper tail

*Significance level

linearity and homoscedasticity in the bivariate distributions. Heteroscedasticity was tested using Bartlett's chi-square procedure, and linearity was evaluated using analysis of variance techniques. None of the TALENT distributions met either of the assumptions, none of the refinery distributions met the homoscedasticity assumption, and only one refinery distribution (Mechanical Comprehension vs. Learning Ability) met the linearity assumption. On the other hand, none of the white collar distributions violated either assumption.

Since corrected correlations are influenced by the form of violations of the assumptions as well as the degree, means and standard deviations of the *X* arrays were examined for systematic trends. The white collar distributions generally appeared to be linear, the refinery distributions tended to flatten in both tails, and the TALENT distributions tended to steepen in the upper tail with one exception (English Total vs. Creativity). Only the TALENT data yielded substantial systematic heteroscedastic trends. Array standard deviations increased with increases in scores on *X* for two of the distributions and decreased for three of the distributions.

Procedure

A computer program was designed (1) to compute the correlation for an observed empirical bivariate data distribution; (2) to truncate the original distribution at successive 10% intervals from the lower end and to compute the corresponding uncorrected and corrected correlations; (3) to modify the original data distribution to control for effects of nonlinearity or heteroscedasticity, or both, by using a pseudorandom number generator to generate new *Y* distributions for the *X* arrays so that the mean of the new array distribution would fall on the regression line and the array standard deviation would equal the standard error of estimate; (4) to compute the correlation for the new score distribution; and (5) to truncate the new distribution at successive 10% intervals and to compute the corresponding uncorrected and corrected correlations.

The corrected correlations were computed using the formula for explicit selection (Gulliksen, 1950, p. 137).

$$r_{xy} = \frac{S_x r_{x*y*}}{(S_x^2 r_{x*y*}^2 + S_{x*}^2 - S_{x*}^2 r_{x*y*}^2)^{\frac{1}{2}}} \tag{1}$$

where X^* and Y^* represent variables for the restricted group. A measure of accuracy was developed by subtracting the Fisher *Z* transformation of the unrestricted correlation from the *Z* transformation of the corrected correlation. The same procedure was also used to estimate the accuracy of the uncorrected correlation.

Results

The data were summarized by computing for each degree of truncation the median error over distributions with similar unrestricted correlations. Median error gives an indication of the extent to which there is a systematic tendency for the corrected correlations systematically to overestimate or to underestimate the unrestricted correlations. This is analogous to bias in estimators computed on random samples. Median error was also computed for estimates using the uncorrected correlations.

Median absolute error was also computed to summarize the magnitude of the errors in estimating the unrestricted correlations. However, the median absolute error was in most instances the same or very nearly the same as median error. Consequently, median absolute error was not reported.

Size of the Unrestricted Correlation

Data on the accuracy of the correction for explicit selection, presented in Tables 2 through 4, show that for small unrestricted correlations between .10 and .25, the corrected correlation was no more accurate than the uncorrected esti-

Table 2
Errors in Estimating Unrestricted Correlations for Moderate Truncation

		10% Truncation		20% Truncation		30% Truncation	
Distribution	r	Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected
Chem. Comp. vs. Learn. Abil.	.678	.065	-.082	.091	-.140	.099	-.201
Mech. Comp. vs. Learn. Abil.	.623	.036	-.096	.053	-.150	.074	-.189
Mech. Comp. vs. Chem. Comp.	.633	.019	-.079	-.005	-.158	-.010	-.202
Inf. I vs. Read. Comp.	.789	.051	-.136	.096	-.190	.144	-.254
Eng. Total vs. Read. Comp.	.738	-.031	-.195	-.026	-.270	-.007	-.340
Median		.036	-.096	.053	-.158	.074	-.202
Appraisal vs. Miller Anal.	.144	-.018	-.031	-.029	-.052	-.044	-.071
Appraisal vs. Nonverb. Reas.	.162	-.003	-.049	-.007	-.067	.008	-.067
Appraisal vs. GZTS Key	.132	.020	-.012	.029	-.020	.046	-.019
Appraisal vs. Biodata Key	.205	.026	-.020	.057	-.015	.046	-.044
Info. I vs. Mature Pers.	.243	.047	.015	.081	.020	.143	.036
Median		.020	-.020	.029	-.020	.046	-.044
English Total vs. Creat.	.512	-.012	-.061	-.040	-.139	-.050	-.173
Math Total vs. Ab. Reas.	.568	.048	-.119	.063	-.156	.071	-.206
Math Total vs. Scien. Int.	.361	.014	-.028	.038	-.036	.072	-.042
Median		.014	-.061	.038	-.139	.071	-.173

Note: Error was computed by subtracting the z equivalent of the population correlation from the z equivalent of the restricted correlation (corrected and uncorrected).

Table 3

Errors in Estimating Unrestricted Correlations for Extreme Truncation

Distribution	r	40% Truncation		50% Truncation		60% Truncation	
		Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected
Chem. Comp. vs. Learn. Abil.	.678	.075	-.255	.097	-.287	.153	-.300
Mech. Comp. vs. Learn. Abil.	.623	.087	-.218	.102	-.252	.113	-.289
Mech. Comp. vs. Chem. Comp.	.633	-.013	-.238	-.057	-.305	-.121	-.378
Inf. I vs. Read. Comp.	.789	-.171	-.294	.207	-.364	.298	-.412
Eng. Total vs. Read. Comp.	.738	.019	-.371	.067	-.416	.127	-.467
Median		.019	-.255	.097	-.305	.127	-.378
Appraisal vs. Miller Anal.	.144	-.052	-.083	-.051	-.087	-.017	-.072
Appraisal vs. Nonverb. Reas.	.162	.012	-.074	-.053	-.120	-.035	-.120
Appraisal vs. GZTS Key	.132	.046	-.026	.056	-.033	.071	-.034
Appraisal vs. Biodata Key	.205	.051	-.051	.068	-.057	.066	-.070
Info. I vs. Mature Pers.	.243	.157	.029	.244	.049	.259	.034
Median		.046	-.051	.056	-.057	.066	-.070
English Total vs. Creat.	.512	-.068	-.214	-.088	-.252	-.092	-.279
Math Total vs. Ab. Reas.	.568	.074	-.258	.074	-.258	.073	-.314
Math Total vs. Scien. Int.	.361	.097	-.060	.138	-.055	.166	-.060
Median		.074	-.214	.074	-.252	.073	-.279

Table 4

Errors in Estimating Unrestricted Correlations for Very Extreme Truncation

Distribution	r	70% Truncation		80% Truncation		90% Truncation	
		Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected
Chem. Comp. vs. Learn. Abil.	.678	.091	-.401	-.060	-.519	-.419	-.684
Mech. Comp. vs. Learn. Abil.	.623	.116	-.344	.029	-.426	-.140	-.519
Mech. Comp. vs. Chem. Comp.	.633	-.194	-.449	-.195	-.474	-.387	-.596
Inf. I vs. Read. Comp.	.789	.351	-.460	.445	-.587	.640	-.720
Eng. Total vs. Read. Comp.	.738	.172	-.506	.299	-.584	.539	-.671
Median		.116	-.449	.029	-.519	-.140	-.671
Appraisal vs. Miller Anal.	.144	.023	.025	.029	.023	.037	-.046
Appraisal vs. Nonverb. Reas.	.162	.058	-.102	.178	-.086	.402	-.101
Appraisal vs. GZTS Key	.132	.093	-.033	-.003	-.086	-.031	-.102
Appraisal vs. Biodata Key	.205	.037	-.098	.086	-.090	.019	-.130
Info. I vs. Mature Pers.	.243	.289	.028	.320	-.005	.471	.003
Median		.058	-.033	.086	-.086	.037	-.101
English Total vs. Creat.	.512	-.074	-.293	-.071	-.318	-.037	-.354
Math Total vs. AB. Reas.	.568	.104	-.356	.225	-.368	.237	-.432
Math Total vs. Scien. Int.	.361	.215	-.077	.225	-.118	.233	-.157
Median		.104	-.293	.225	-.318	.233	-.354

mate. This was true over all degrees of truncation. The main difference between corrected and uncorrected correlations was that the uncorrected value consistently tended to underestimate the unrestricted correlation, while there was a slight tendency for the corrected value to overestimate the unrestricted correlation.

For large unrestricted correlations in the range .60 to .80, the corrected correlation was always a more accurate estimate of the unrestricted correlation than the uncorrected correlation; and the greater the degree of the truncation, the greater the error occasioned by using the uncorrected estimate as compared with the corrected value. The uncorrected correlation consistently underestimated the unrestricted correlation, while the corrected correlation tended to overestimate the unrestricted correlation. This overestimation probably occurred because most of these distributions were flat in the lower tail—the exception being the distribution of English Total versus Reading Comprehension, which was slightly steeper in both tails. This latter condition, coupled with smaller array variances in the upper tail, led to underestimation for small degrees of truncation and overestimation for large degrees of truncation.

Unrestricted correlations in the medium range .30 to .55 behaved like the high unrestricted correlations. With few exceptions, the corrected correlation was a more accurate estimate of the unrestricted correlation than was the uncorrected value.

Degree of Truncation

The data showed that the correction became progressively less accurate with increasing truncation. Thus, for 30% or less truncation, the median error was around .05 correlation points for medium and high unrestricted correlations and about .03 for low unrestricted correlations. However, for high unrestricted correlations, the relative error due to not correcting is larger for higher degrees of truncation. Thus, in one sense it is more important to use corrected estimates of the unrestricted correlations under conditions of high truncation than under low truncation.

Violations of Assumptions

Inspection of Tables 2 through 4 shows no clear pattern of results due to violations of the assumptions of homoscedasticity and linearity. The four distributions that did not show statistically significant departures from either homoscedasticity or linearity were also distributions with low unrestricted correlations. For these distributions, the corrected correlation was no more accurate than the uncorrected correlation as an estimate of the unrestricted correlation. It would appear that the inaccuracy of the correction for low unrestricted correlations cannot be attributed to excessive violations of assumptions.

To explore more deeply the effect of violations of assumptions, the 13 bivariate distributions were modified in three ways: (1) the array standard deviations were equalized, removing any effects of heteroscedasticity; (2) the slope of the regression line was controlled to a constant value (by adjusting array means), thus removing any effects of nonlinearity; and (3) both heteroscedasticity and nonlinearity were removed. The results for 50% truncation are presented in Table 5. These data show that when array variances were equalized, thereby homogenizing the empirical distributions, there was little effect on the accuracy of the correction. On the other hand, when the empirical distributions were linearized, accuracy was greatly improved; and, of course, when the empirical distributions were normalized, the accuracy of the correction was best. These data indicate that the correction is very sensitive to departures from linearity but is relatively insensitive to heteroscedasticity, even where there is a consistent trend in the array standard deviations.

Conclusions and Discussion

This study examined 13 empirical bivariate distributions with respect to accuracy of restric-

Table 5

Effect of Violation of Assumptions on Corrected Correlations

Distribution	r	Uncorrected	Corrected			
			Original	Homogenized	Linearized	Normalized
Chem Comp. vs. Learn. Abil.	.678	-.287	.097	.096	-.007	-.020
Mech. Comp. vs. Learn. Abil.	.623	-.252	.102	.098	.012	.016
Mech. Comp. vs. Chem. Comp.	.633	-.305	-.057	-.050	.034	.018
Inf. I vs. Read. Comp.	.789	-.364	.207	.230	-.022	.003
Eng. Total vs. Read. Comp.	.738	-.416	.067	-.027	.061	-.008
Median		-.305	.097	.096	.012	.003
Appraisal vs. Miller Anal.	.144	-.087	-.051	-.070	-.012	-.008
Appraisal vs. Nonverb Reas.	.162	-.120	-.035	-.071	-.031	-.035
Appraisal vs. GZTS Key	.132	-.033	.056	.063	.006	.001
Appraisal vs. Biodata Key	.205	-.057	.068	.054	-.021	-.023
Info. I vs. Mature Pers.	.243	.049	.244	.230	.001	-.002
Median		-.057	.056	.054	-.012	-.008
English Total vs. Creat.	.512	-.252	-.088	-.143	.063	-.004
Math Total vs. Ab. Reas.	.568	-.258	.074	.133	-.089	-.021
Math Total vs. Scien. Int.	.361	-.055	.138	.201	-.036	.009
Median		-.252	.074	.133	-.036	-.004

Note: Homogenized means that original distribution was modified to equalize the array variances while keeping other features the same. Linearized means that original distribution was modified so that array means fell exactly on the regression line. Normalized means that both modifications were made to the original distribution.

tion in range formulas for explicit selection. The following conclusions were supported by the data.

1. The corrected sample correlation was no more accurate than the uncorrected sample correlation for low unrestricted population correlations in the range from .10 to .25. This was true for distributions in which there were no statistically significant violations of either linearity or homoscedasticity.
2. For large unrestricted population correlations in the range .60 to .80, the corrected sample correlation was always more accurate than the uncorrected sample correlation. While the corrected sample correlations were less accurate for large degrees of truncation compared with moderate truncation, the difference between the corrected estimate and the uncorrected estimate tended to increase progressively with larger amounts of truncation. Thus, in estimating the unrestricted correlation, it is more important in terms of reducing relative error to use the corrected estimate when truncation is extreme.
3. For moderate unrestricted population correlations in the range .30 to .55, the corrected estimates were more accurate than the uncorrected sample correlations.
4. The correction was very sensitive to moderate departures from linearity but was quite robust in the face of rather substantial departures from homoscedasticity. This was true even when there was a systematic trend in the array variances.

Novick and Thayer (1969), in their empirical study of the correction for restriction in range due to explicit selection, attempted to improve the correction by relaxing the homoscedasticity assumption. In two attempts they found that relaxation of the homoscedasticity assumption had little effect on the accuracy of the correction. Results from the present study suggest that since departure from homoscedasticity has little effect on the accuracy of the correction, relaxation of the homoscedasticity assumption will not be helpful in improving the correction.

The finding that the correction is sensitive to moderate departures from linearity poses a serious problem for researchers, especially when the unrestricted correlation is low. There were no typical departures from linearity that were consistent across the distributions studied. Consequently, it appears that the corrected correlation may systematically overestimate or underestimate the unrestricted correlation, depending on ordinarily unknown characteristics of the population distribution under study.

References

American Institutes for Research. *The Project TALENT data bank: A handbook.* Palo Alto, CA: American Institutes for Research, 1972.

Forsyth, R. A. An empirical note on correlation coefficients corrected for restriction in range. *Educational and Psychological Measurement,* 1971. *31,* 115–123.

Greener, J. M., & Osburn, H. G. *Accuracy of corrections for restriction in range due to explicit selection in heteroscedastic and nonlinear distributions.* Unpublished manuscript, 1977. (Available from [Department of Psychology, University of Houston, Houston, TX 77004]).

Gulliksen, H. *Theory of mental tests.* New York: Wiley, 1950.

Lawley, D. A note on Karl Pearson's selection formulae. *Royal Society of Edinburgh, Proceedings, Section A,* 1943–1944, *62,* 28–30.

Lord, F. M., & Novick, M. R. *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley, 1968.

Meredith, W. M. The estimation of criterion parameters from a biased sample (Doctoral Dissertation, University of Washington, 1958). *Dissertation Abstracts,* 1958, *19,* 871.

Novick, M. R., & Theyer, D. T. *An investigation of the accuracy of the Pearson selection formulas* (Research Memorandum RM–69–22). Princeton, NJ: Educational Testing Service, 1969.

Pearson, K. Mathematical contributions to the theory of evolution-XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society, London, Series A*, 1903, *200*, 1–66.

Rydberg, S. *Bias in prediction*. Stockholm: Almqvist & Wiksell, 1963.

Srinivasan, V., & Weinstein, A. G. Effects of curtailment on an admissions model for a graduate management program. *Journal of Applied Psychology*, 1973, *58*, 339–346.

Acknowledgement

Acknowledgement for use of the TALENT data is given to American Institutes for Research, which maintains the project TALENT data bank.

Author's Address

H. G. Osburn, Department of Psychology, University of Houston, Houston, TX 77004.