Journal of Industrial Psychology, 1996, 22(1), 19-27
Tydskrif vir Bedryfsielkunde, 1996 22(1), 19-27

THE DEVELOPMENT OF A STATISTICAL PROCEDURE TO CORRECT THE EFFECTS OF RESTRICTION OF RANGE ON VALIDITY COEFFICIENTS

J M SCHEPERS

Department of Human Resource Management
Rand Afrikaans University

OPSOMMING

In die geldigheidsbepaling van toetse wat vir keuringsdoeleindes gebruik word, is die verkreë geldigheidskoëffisiënte sonder uitsondering onderskattings van die ware geldighede as gevolg van eksplisiete en implisiete keuring ten opsigte van die tersaaklike veranderlikes. Sowel eksplisiete as implisiete keuring lei tot inperking van die variasiewydte van die relevante veranderlikes, en dit reduseer om die beurt weer die verkreë geldighede. 'n Formele bewys hiervoor word in die referaat gegee. 'n Aantal navorsers het formules ontwikkel om *steekproefgeldighede* te korrigeer ten einde beter beramings van die ware geldighede te verkry (Pearson, 1903; Thorndike, 1949: Gulliksen, 1950; Rygberg, 1962 en Lord & Novick, 1968). Dit is egter bykans onmoontlik om op hierdie wyse 'n geheelbeeld van die probleem van inperking van variasiewydte te vorm. In die huidige referaat is 'n ander benadering gevolg: Populasiekorrelasies is bereken vir verskillende grade van afknotting van die eksplisiete keuringsveranderlike. Dit is gedoen vir populasiekorrelasies wat wissel van 0,10 tot 0,99. Voorts is 'n grafiese voorstelling voorberei, wat die krimpring van die populasiekorrelasies vir verskillende afknottingsverhoudings, uitbeeld.

ABSTRACT

In the validation of tests used for selection purposes, the obtained validity coefficients are invariably underestimates of the true validities, due to *explicit* and *implicit* selection in respect of the relevant variables. Both explicit and implicit selection leads to restriction of range of the relevant variables, and this in turn reduces the obtained validites. A formal proof for this is given. A number of researchers have developed formulae for correcting *sample validities* in order to get better estimates of the true validities (Pearson, 1903; Thorndike, 1949; Gulliksen, 1950; Rydberg, 1962 and Lord & Novick, 1968). It is, however, virtually impossible to obtain a complete view of the problem of restriction of range in this way. In the present paper a different approach has been followed: Population correlations have been computed for various degrees of truncation of the explicit selection variable. This has been done for population correlations ranging from 0,10 to 0,99. A graphical display, indicating the *shrinkage* of the population correlations for various truncation ratios, has been prepared.

In validating tests for use in personnel selection, special care should be taken to ensure that the obtained validities are neither underestimates nor overestimates of the true validities. This calls for a thorough study of all the factors, both psychological and statistical, that impinge on the validation process. In particular the following factors or conditions merit special attention in this regard:

1. The degree of *natural* or *self-selection* that has taken place prior to the onset of the study;
2. the degree of *direct* or *indirect selection* that has taken place prior to the *maturation* of the criterion measures;
3. the amount of *learning* or *growth* that has taken place prior to the maturation of the criterion measures;
4. the metric properties of both the predictors and the criteria;
5. the nature of any *transformations* that might have been applied to the test and/or criterion scores, and
6. indications of *test bias*, particularly cultural bias.

Next, a brief account will be given of each of the factors or conditions mentioned above:

A host of factors are constantly at work in nature, shaping the attributes of man. Through *natural selection* the *variability* of certain attributes are *restricted*, whilst the variability of other attributes are increased. A typical case in point concerns the intelligence of children at school. At primary school level the distribution of IQ's is very wide, i.e. the primary school population is very heterogeneous with respect to intelligence. By contrast the high school population is much more homogeneous. Similarly, the population of university students is restricted towards the upper-end of the continuum of intelligence. The selection mechanism is obvious – students without the necessary abilities drop out of the school system at various stages. However, in South Africa the picture is much more complex. During the years of "apartheid" gross *inequalities* existed in the various educational systems in South Africa. The quality of education afforded Black children were generally regarded as inferior to those of White children. A large proportion of Black children never attended school, and many of those who did attend, left school *prematurely*. The formative value of formal schooling was therefore quite different for Whites and Blacks. Natural selection might therefore have resulted in quite different outcomes for the various *cultural groups* in South Africa.

Coupled with the process of natural selection, there is another selection mechanism at work, which could be called *self-selection*: Persons with common interests flock together and establish interest groups, study particular subjects, and prepare themselves for particular jobs or professions.

Pearson (1903) was one of the first statisticians to recognise the effects of *natural selection* on the *variability* of particular attributes of organisms, and how that influenced the correlation between the attributes.

Direct or explicit selection can best be explained with reference to a real selection situation. Suppose a particular university

Requests for reprints should be addressed to J. M. Schepers, Department of Human Resource Management, Rand Afrikaans University, P.O. Box 524, Auckland Park 2006

decided to accept all applicants with IQ's of 115 and higher, and to reject all applicants with IQ's of 114 and lower. Then the first step would be to test all applicants with a suitable test of intelligence, and if their IQ's are 115 or higher they will be accepted for enrolment, otherwise they will be rejected. In this case there is *direct or explicit selection* in terms of their scores in the test of intelligence. Next, the achievements of the students that were accepted for training are evaluated in terms of their scores in the final examination at the end of the academic year. If the test of intelligence is statistically significantly correlated with the examination results, the *variability* of the examination results will be *restricted*, depending on the magnitude of the correlation. I the correlation is high the restriction of variability will be severe. This is known as *indirect* or implicit selection.

The amount of learning or growth that has taken place prior to the maturation of the criterion measures, is of critical importance in validation studies. As will be shown later, validity coefficients *shrink* if the *variance* of the criterion measure is restricted due to explicit selection on the predictor. However, there is often a concomitant change in the variance of the criterion due to *learning.*

Schepers (1962), using a complex psychomotor learning task, found that "the variances of the performance scores in respect of the various samples used, reveal a very interesting gradient. The variances of the first trial are generally low, compared with subsequent trials. This would seem to suggest that the responses made during the first trial are mainly exploratory in nature, and serve to orientate the subject. Following the first trial, the variances tend to increase first (till fifth or sixth trial) and then to decrease again". By the end of the sixth trial most of the subjects had learnt the combination of lights and controls. "This period would seem to coincide with the stage when the group is most variable. As learning continues, and the task becomes overlearnt, the variability of the scores decrease again" (p. 309).

From the foregoing it should be obvious that the stage at which the criterion scores are collected is of critical importance in validation studies. The stage when *variability* is at its *maximum* would seem to be the ideal stage for collecting the criterion data.

Alexander (1988) studied the effects of *range restriction* and *range enhancement* on correlations, and came to the conclusion that "increasing the variance on one or both variables in an intact group has no necessary influence on the correlation" (p. 774). To support his argument, he uses the following illustration: A group is measured at Time 1 on a predictor (X) and on a job performance variable (Y). They are then subjected to a *training program* that results in every individual's job performance being multiplied by a constant (c) at Time 2. He then correctly points out that this is a *linear transformation* of the criterion scores, and that the correlation between any two variables is *invariant* under such linear transformations. His final conclusion, however, is wrong, largely because of a misconception of the process of learning.

In order to illustrate the nature of his misconception, it is necessary to take a critical look at the formula for the *reliability of difference scores* (cf. Schepers, 1992, pp. 53-57). The formula for the reliability of such scores can be written as follows:

$$\rho_{vv'} = \frac{(\overline{\rho} - \rho_{xy})}{(1 - \rho_{xy})} \qquad \ldots\ldots \quad 1.1$$

where $\overline{\rho}$ = the *average* of the reliabilities of the pre-test scores and the post-test scores, and

ρ_{xy} = the correlation between the pre-test and post-test scores.

From equation 1.1 it is clear that there are two conditions for a *high* reliability of difference scores, namely

1. A *high* average reliability of the pre- and post-test scores, and
2. a *low* correlation between the pre- and post-test scores.

The *higher* the correlation between the pre- and post-test scores, the *lower* is the reliability of the difference scores:

If the reliabilities of the pre- and post-test scores are equal, i.e.

$\rho_{xx'} = \rho_{yy'}$, then

$$\rho_{vv'} = \frac{(\rho_{xx'} - \rho_{xy})}{(1 - \rho_{xy})} \qquad \ldots\ldots \quad 1.2$$

The *maximum* value of ρ_{xy} is equal to the *index of reliability* of the test, and is given by $\sqrt{\rho_{xx'}}$.

Substitution of $\sqrt{\rho_{xx'}}$ for ρ_{xy} in equation 1.2 gives:

$$\rho_{vv'} = \frac{\rho_{xx'} - \sqrt{\rho_{xx'}}}{1 - \sqrt{\rho_{xx'}}} \qquad \ldots\ldots \quad 1.3$$

$$= \frac{\sqrt{\rho_{xx'}}(\sqrt{\rho_{xx'}} - 1)}{-(\sqrt{\rho_{xx'}} - 1)} \qquad \ldots\ldots \quad 1.4$$

$$= -\sqrt{\rho_{xx'}} \qquad \ldots\ldots \quad 1.5$$

In this case the reliability of the difference scores is *extremely low*. This is actually the case that holds if each subject's score is *multiplied by a constant.* The ranks of the subjects have not changed at all.

On the other hand, the *lower* the correlation between the pre- and post-test scores, the *higher* the reliability of the difference scores:

If $\rho_{xy} = 0$, then

$$\rho_{vv'} = \frac{\rho_{xx'} - 0}{1 - 0} \qquad \ldots\ldots \quad 1.6$$

$$= \rho_{xx'}$$

A *low* correlation between the pre- and post-test scores implies *differential growth* on the part of the test subjects as a function of the intervention to which they have been subjected. There is thus a *radical change* in the ranks of the test subjects from the pre-test to the post-test situation.

From the foregoing it should be apparent that if real learning had taken place, the *criterion variance* might well have been *enhanced,* and that would certainly influence the test-criterion correlation. Alexander's (1988, p. 774) conclusion is therefore not justified.

In validating a test against an external criterion, it is important to ensure that both variables yield normative rather than ipsative scores, and that both measures are optimally reliable.

Ipsative scores can be used to determine the relative strengths of intra-individual differences, but not inter-individual differences, with respect to psychological attributes. Clemans (1956, p. 52) sounds a warning in this regard: "Ipsative scores are relative scores. It is quite possible that a person obtaining a low ipsative score on a particular trait actually possesses more of the characteristic in question than a person obtaining a much higher ipsative score. It is imperative that users of ipsative variables interpret them in the relative sense only".

As far as reliability is concerned, it is obvious that the best possible test cannot predict a totally *unreliable* criterion. A proper criterion, with acceptable metric properties, should therefore first be constructed. Corrections for *attenuation* can then be applied to the criterion, and for certain purposes, also to the test.

It is generally known that linear transformations of either or both variables involved in a correlation, leaves the correlation *invariant*, i.e. unchanged. But this is not true of *non-linear transformations* such as are involved in the *normalisation* of test scores. The effects of normalisation are difficult to predict: It will raise certain correlations, and reduce others. However, it is good practice to normalise variables wherever possible. This also applies to all the variables involved in validation studies.

The issue of *test bias* is very complex and cannot be dealt with fully in this context. However, certain aspects of test bias that are related to the central theme of this paper, will be dealt with here.

To start off with an appropriate definition of test bias is called for. For our purpose the definition of Cleary (1968, p. 115) will suffice: "A test is biased for members of a subgroup of the population if, in the prediction of a criterion for which the test was designed, consistent nonzero errors of prediction are made for members of the sub-group". This implies that in the prediction of the criterion, the test consistently over-predicts or under-predicts for a particular subgroup of the population. Statistically, this implies that there is not a *common regression* line for the various subgroups of the population. As will become apparent later, when the assumptions underlying corrections for restriction of range are discussed, test bias leads to a *violation* of the *basic assumptions* underlying these corrections.

A common regression line for all the subgroups of the population is not a guarantee that there is no bias in the regression system. If the criterion is *multidimensional*, the test could correlate with different weighted composites thereof for the various subgroups. A careful study of the criterion is therefore essential.

From the foregoing it is clear that there are several conditions that lead to a restriction of range of the predictor, and indirectly also of the criterion. Learning or growth typically leads to *enhancement* of the variance of the criterion. *Attenuation* due to the *unreliability* of both the predictor and the criterion, leads to an *underestimate* of the correlation between the two variables. Correcting a validity coefficient for attenuation of the criterion, seems justified, as the best predictor cannot predict a totally unreliable criterion.

As was intimated earlier on, restriction of range invariably leads to a shrinkage of the obtained correlation. Accordingly a brief proof of this will be given:

Starting off with the regression of Y on X, the *coefficient of determination* (ρ^2_{xy}) can be derived as follows:

$$\hat{Y} = \rho_{xy}\frac{\sigma_y}{\sigma_x}(X_i - \mu_x) + \mu_y \qquad \ldots\ldots \quad 2.1$$

$$(\hat{Y} - \mu_y) = \rho_{xy}\frac{\sigma_y}{\sigma_x}(X_i - \mu_x) \qquad \ldots\ldots \quad 2.2$$

$$(\hat{Y} - \mu_y)^2 = \rho^2_{xy}\frac{\sigma^2_y}{\sigma^2_x}(X_i - \mu_x)^2 \qquad \ldots\ldots \quad 2.3$$

$$\sum_{i=i}^{N}(\hat{Y} - \mu_y)^2 = \rho^2_{xy}\frac{\sigma^2_y}{\sigma^2_x}\sum_{i=i}^{N}(X_i - \mu_x)^2 \qquad \ldots\ldots \quad 2.4$$

$$\therefore\ \rho^2_{xy} = \frac{\sigma^2_x}{\sigma^2_y}\left[\frac{\sum_{i=i}^{N}(\hat{Y} - \mu_y)^2}{\sum_{i=i}^{N}(X_i - \mu_x)^2}\right] \qquad \ldots\ldots \quad 2.5$$

$$= \left[\frac{\sum(X_i - \mu_x)^2/N}{\sum(Y_i - \mu_y)^2/N}\right]\cdot\left[\frac{\sum(\hat{Y} - \mu_y)^2}{\sum(X_i - \mu_x)^2}\right] \ldots \quad 2.6$$

$$= \frac{\sum(\hat{Y} - \mu_y)^2}{N}\Big/\frac{\sum(Y_i - \mu_y)^2}{N} \qquad \ldots\ldots \quad 2.7$$

$$= \frac{\sigma^2_{\hat{Y}}}{\sigma^2_y} \qquad \ldots\ldots \quad 2.8$$

From equation 2.8 it is clear that the coefficient of determination is equal to the *ratio* of the predicted variance $[\sigma^2_{\hat{Y}}]$ relative to the variance of the criterion $[\sigma^2_y]$.

Starting off with the standard error of estimate, the *residual variance* $[\sigma^2_{x.y}]$ can be derived as follows:

$$\sigma_{y.x} = \sigma_y\sqrt{1 - \rho^2_{xy}} \qquad \ldots\ldots \quad 2.9$$

Squaring both sides:

$$\sigma^2_{y.x} = \sigma^2_y(1 - \rho^2_{xy}) \qquad \ldots\ldots \quad 2.10$$

$$\therefore\ (1 - \rho^2_{xy}) = \frac{\sigma^2_{y.x}}{\sigma^2_y} \qquad \ldots\ldots \quad 2.11$$

$$\text{And}\ \ \rho^2_{xy} = 1 - \frac{\sigma^2_{y.x}}{\sigma^2_y} \qquad \ldots\ldots \quad 2.12$$

From equations 2.8 and 2.12 it follows that

$$\frac{\sigma^2_{\hat{Y}}}{\sigma^2_y} = 1 - \frac{\sigma^2_{y.x}}{\sigma^2_y} \qquad \ldots\ldots \quad 2.13$$

$$\text{i.e.}\ \ \frac{\sigma^2_{\hat{Y}}}{\sigma^2_y} = \frac{\sigma^2_y - \sigma^2_{y.x}}{\sigma^2_y} \qquad \ldots\ldots \quad 2.14$$

$$\therefore\ \ \sigma^2_{\hat{Y}} = \sigma^2_y - \sigma^2_{y.x} \qquad \ldots\ldots \quad 2.15$$

The coefficient of determination can therefore be written as follows:

$$\rho_{xy}^2 = \frac{\sigma_{\hat{Y}}^2}{\sigma_y^2} \quad \text{..... } 2.16$$

$$= \frac{\sigma_y^2 - \sigma_{y.x}^2}{\sigma_y^2} \quad \text{..... } 2.17$$

$$= 1 - \frac{\sigma_{y.x}^2}{\sigma_y^2} \quad \text{..... } 2.18$$

In order to interpret equation 2.18 correctly, it is important to realise that it rests on two *basic assumptions*, viz. *linearity of the regression* of Y on X, and *homoscedasticity* of the bivariate distribution (YX). In dealing with a *population*, it is assumed that these assumptions hold throughout the domain of X. If equality of the variances of the arrays of Y, conditional on X, holds, then the residual variance of Y is *constant* throughout the domain of X (Lord & Novick, 1968, pp. 142-143; Howell, 1992, p. 243 and Magnusson, 1967, p. 145).

In the light of the foregoing it is clear that the coefficient of determination (ρ_{xy}^2) will shrink if the variance of the criterion is restricted due to explicit selection on X. Restriction of range will therefore lead to a shrinkage in the obtained correlation.

From the foregoing it is clear that restriction of range invariably leads to a *depression* of validity coefficients. It is therefore necessary to develop an appropriate procedure to correct the effects of restriction of range on correlation coefficients. This then is the primary objective of the present paper.

Pearson (1903) was the first to develop a formula for correcting correlations for restriction of range. Following in his footsteps, Thorndike (1949) and Gulliksen (1950) developed slight variants of Pearson's formula, based on the *same basic assumptions*. Lord and Novick (1968) followed the same route as Gulliksen and developed a formula (see equation 6.8.5, p. 143) which is formally identical to Gulliksen's Formula 18 (cf. Gulliksen, 1950, p. 137), but expressed in a different form. Rydberg (1962), starting off with the same basic assumptions as his predecessors, developed methods of correcting correlations for indirect restriction of range with *non-interval* data.

Mendoza and Mumford (1987) looked into the joint effects of attenuation and range restriction "(a) when range restriction results from direct restriction on the predictor, and (b) when range restriction results from direct restriction on a latent variable defined by the predictor" (p. 282).

Their contribution is particularly useful in cases where *restriction on ability* has taken place, rather than on the *predictor variable* itself. Validation studies involving *test-retest designs* would fall in this category.

The fact that their correction formulae (equations (5) and (10)) correct for both *restriction of range* and *attenuation* is a mixed blessing. In evaluating a selection device, the *unreliability* of the *criterion* is relevant, but not the unreliability of the predictor itself. After all, the goal is to evaluate the predictor as it *is*, not as it *might be*. However, this problem can easily be rectified.

All of the corrections for restriction of range, dealt with so far, rest on the same basic assumptions, viz. *linearity* of the *regression* of Y on X, and *homoscedasticity* of the bivariate distribution (XY). However, Gross (1982, p. 798) has shown that the correction formula can hold even if both assumptions are violated.

He has shown that a *sufficient condition* for the correction formula to hold, is that

$Q = \sqrt{S_e^2/s_e^2}\ /(B/b) = 1$, where

S_e^2 = error variance of total group;
s_e^2 = error variance of selected group;
B = slope of regression line of total group, and
b = slope of regression line of selected group.

If $Q > 1$, and both B and b are *positive*, the correction formula will *overestimate* the true validity. And if $Q < 1$, the correction formula will *underestimate* the true validity (p. 799).

This contribution of Gross is of great practical value and ought to be routinely applied whenever corrections for restriction of range are made. In practice it will certainly facilitate the interpretation of research findings.

Gross and Kagen (1983) asked the question as to whether it is always advantageous to correct for restriction of range. Using a Monte Carlo simulation they calculated the *expected values* for uncorrected and corrected *correlations* as a function of different group sizes, population correlations and selection ratios. They found that both the corrected and uncorrected correlations are *negatively biased estimators* of the population correlation, but that in every case the *uncorrected correlations* were *more biased* than the corrected correlations (p. 393).

They also calculated the *expected mean square errors* for the uncorrected and corrected correlations as a function of different group sizes, population correlations and selection ratios. They found that for *low values* of the population correlation ($\rho_{xy} = 0,30$) the uncorrected correlation has a *smaller* expected mean square value in 21 of 27 cases, but for higher population correlations the advantage quickly disappears (p. 393). In general the uncorrected correlations proved advantageous for *extreme selection ratios*.

The findings of Gross and Kagen are somewhat surprising as very low correlations hardly shrink, even under extreme selection. It would be interesting to see the corresponding Q-values for these cases.

House (1983), using a formula to correct for restriction of range, calculated both the corrected and uncorrected predictive validity of the Graduate Record Examination. He found that the uncorrected correlation was 0,15, and the corrected correlation 0,17. He finally comes to the conclusion that "there is no evidence that improvement in a correlation corrected for restriction is dependent upon the magnitude of the original predictive validity" (p. 710). As will be shown in the main study, this conclusion is completely fallacious. Both very high and very low correlations are only slightly influenced by restriction of range, but the middle range of correlations (0,4 to 0,8) are strongly influenced.

Boone and Lewis (1980) were interested in the effects of *recruitment* on corrections for restriction of range. They used several different unrestricted *applicant groups*, and a selected group. The *predictor variances* of the various unrestricted and selected groups were quite different. Using Gulliksen's (1950, p. 137) Formula 18, they calculated the corrected correlations for each applicant group, with the selected group as a constant. "They found that as the values of the correlation (R_{xy}) move towards the middle values, the discrepancies between the estimated validity coefficients for the different S_x/SS_x ratios become even more pronounced" (p. 929).

From the findings of Boone and Lewis (1980) it is clear that different recruitment philosophies have different effects on the *variances* of the *abilities* of candidates that apply for a particular position. If the advertisement is *very specific* regarding the qualifications of applicants, a greater restriction of range will take place than if the advertisement is *fairly general*.

From the foregoing it should be clear that restriction of range plays an important role in the validation of tests, and that a variety of formulae have been developed to correct the effects of restriction of range. However, the users of these correction formulae are not always convinced that the corrections are *justified*: Some researchers live under the misconception that the *magnitude* of the corrections are generally small and do not justify the effort of making the corrections. Others are concerned about *violations* of the *basic assumptions* underlying the correction formulae. It is therefore clear that further research on the topic is justified.

In the present paper a different approach will be followed – instead of correcting sample correlations, *population correlations* will be calculated for various *degrees of truncation* of the predictor variable. This should have the following advantages:

* The *assumptions* underlying the correction formula will be met in the *universe*;
* a set of *statistical tables* can be prepared which will facilitate future calculations, and
* a complete *view* can be obtained of the *magnitudes* of the corrections for different population correlations, and for different degrees of truncation of the predictor variable.

METHOD

Firstly, a computational formula will be derived to calculate *population correlations* for various degrees of truncation of the predictor variable.

Secondly, *population variances* will be calculated for various degrees of restriction of range.

Finally, a set of *statistical tables* will be prepared, giving population correlations for various degrees of restriction of range of the predictor variable. These correlations will also be *graphically displayed* so as to give an overview of the *magnitudes* of the corrections.

As mentioned earlier in this paper, all the correction formulae developed thus far, rest on two basic assumptions, viz. *linearity of the regression* of Y on X, throughout the domain of X, and *homoscedasticity* of the bivariate distribution (XY). The *linearity assumption* implies that the *slope* of the regression line of the *selected group* is equal to the *slope* of the regression line of the *unselected group*, i.e. that the two regression lines are *parallel* to one another. The *homoscedasticity assumption* implies *equality* of the variances of the arrays of Y, conditional on X. This in turn implies that the residual variance (error variance) of Y is *constant* throughout the domain of X (Lord and Novick, 1968, pp. 142-143). The essentials of these two assumptions are graphically depicted in Figures 1 and 2, respectively.

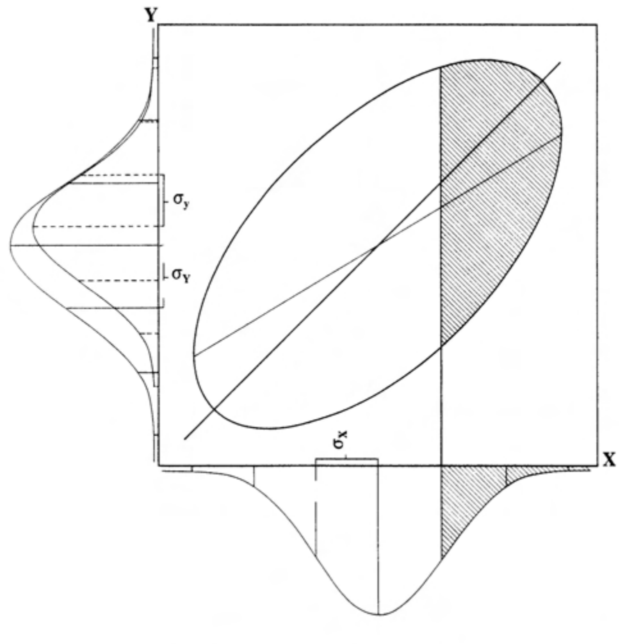

Figure 1: Bivariate normal distribution, indicating the effects of explicit selection on X.

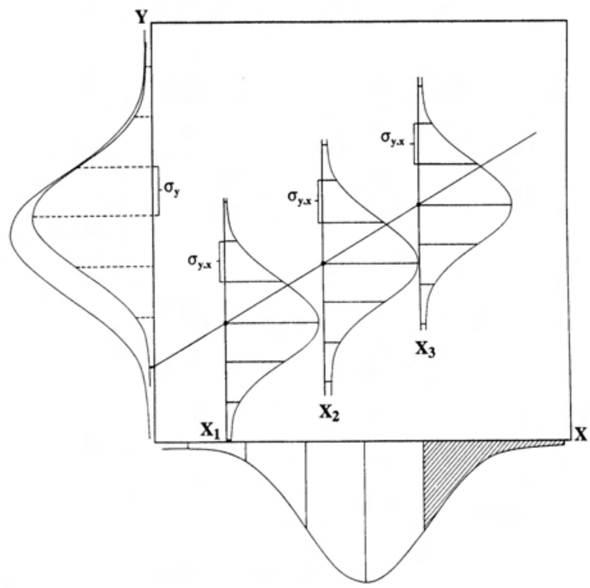

Figure 2: Regression line, indicating the residual variance of Y, conditional on X.

If lower case subscripts are used to refer to the selected group, and upper case subscripts to refer to the unselected group, we have:

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = \rho_{XY}\frac{\sigma_Y}{\sigma_X} \qquad \ldots\ldots \quad 3.1$$

and $$\sigma_y\sqrt{1-\rho_{xy}^2} = \sigma_Y\sqrt{1-\rho_{XY}^2}, \qquad \ldots\ldots \quad 3.2$$

where equation 3.1 states that the *slope* of the regression line of the *selected group* is equal to the slope of the regression line of the *unselected group*. And equation 3.2 states that the *standard error of estimate* of the selected group is equal to the standard error of estimate of the unselected group.

Next, a computational formula will be developed to calculate population correlations for various degrees of truncation of the predictor variable:

From equation 3.1 it follows that

$$\sigma_Y = \frac{\rho_{xy}\sigma_y\sigma_X}{\rho_{XY}\sigma_x} \quad \ldots\ldots \quad 3.3$$

From equation 3.2 it follows that

$$\sigma_y^2(1-\rho_{xy}^2) = \sigma_Y^2(1-\rho_{XY}^2) \quad \ldots\ldots \quad 3.4$$

Substituting for σ_Y^2 in equation 3.4 we have:

$$\sigma_y^2(1-\rho_{xy}^2) = \frac{\rho_{xy}^2\sigma_y^2\sigma_X^2}{\rho_{XY}^2\sigma_x^2}(1-\rho_{XY}^2) \quad \ldots\ldots \quad 3.5$$

Dividing both sides by σ_y^2:

$$(1-\rho_{xy}^2) = \frac{\rho_{xy}^2\sigma_X^2}{\rho_{XY}^2\sigma_x^2}(1-\rho_{XY}^2) \quad \ldots\ldots \quad 3.6$$

$$\frac{(1-\rho_{xy}^2)}{\rho_{xy}^2} = \frac{(1-\rho_{XY}^2)\sigma_X^2}{\rho_{XY}^2\sigma_x^2} \quad \ldots\ldots \quad 3.7$$

$$\left[\frac{1}{\rho_{xy}^2} - 1\right] = \frac{\sigma_X^2}{\sigma_x^2}\left[\frac{1-\rho_{XY}^2}{\rho_{XY}^2}\right] \quad \ldots\ldots \quad 3.8$$

$$\frac{1}{\rho_{xy}^2} = \frac{\sigma_X^2 - \sigma_X^2\,\rho_{XY}^2}{\sigma_x^2\,\rho_{XY}^2} + 1 \quad \ldots\ldots \quad 3.9$$

$$= \frac{\sigma_X^2 - \sigma_X^2\,\rho_{XY}^2 + \sigma_x^2\,\rho_{XY}^2}{\sigma_x^2\,\rho_{XY}^2} \quad \ldots\ldots \quad 3.10$$

$$\therefore \rho_{xy}^2 = \frac{\sigma_x^2\,\rho_{XY}^2}{\sigma_X^2 - \sigma_X^2\sigma_{XY}^2 + \sigma_x^2\,\rho_{XY}^2} \quad \ldots\ldots \quad 3.11$$

$$\therefore \rho_{xy} = \frac{\sigma_x\,\rho_{XY}}{\sqrt{\sigma_X^2 - \sigma_X^2\,\rho_{XY}^2 + \sigma_x^2\,\rho_{XY}^2}} \quad \ldots\ldots \quad 3.12$$

Without loss in generality the population variance (σ_X^2) can be set equal to *unity*:

$$\rho_{xy} = \frac{\sigma_x\,\rho_{XY}}{\sqrt{1-\rho_{XY}^2(1-\sigma_x^2)}} \quad \ldots\ldots \quad 3.13$$

Equation 3.13 can be used to calculate population correlations for different degrees of truncation of the predictor variable, provided the corresponding *variances* of the predictor variable are available.

The next step in the calculation of population correlations is therefore to compute *population variances* for different truncation ratios (p) of the predictor variable. However, before this is done certain concepts need to be clarified:

Let the normal distribution, depicted in Figure 3, represent the predictor variable used in the selection process. Then p would indicate the proportion of candidates that is rejected, and 1 – p would indicate the proportion that is accepted. The standardised point of truncation is designated by (k).

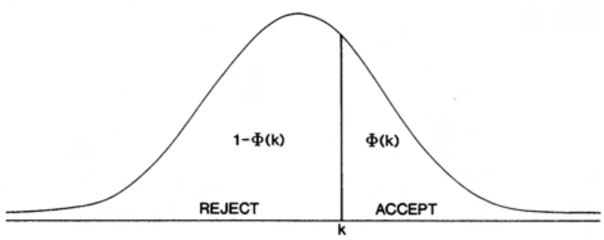

Figure 3: Normal distribution, indicating the proportion of candidates accepted and rejected.

The area under the normal curve, designated Φ(k) (see Figure 3), is given by the normal integral

$$\Phi(\mathrm{k}) = \frac{1}{\sqrt{2\pi}}\int_{\mathrm{k}}^{\infty} e^{\frac{-x^2}{2}}dx \quad \ldots\ldots \quad 4.1$$

and the density is given by the standard normal density function

$$\phi(x) = \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} \quad \ldots\ldots \quad 4.2$$

Dividing the density function by Φ(k) will set the area under the curve equal to *unity*.

$$\mathrm{f}(x) = \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})} \quad \ldots\ldots \quad 4.3$$

The *mean* of the distribution (μ) is given by:

$$\mu = \int_k^{\infty} x.f(x)\,dx \quad \ldots\ldots \quad 4.4$$

$$\mu = \int_k^{\infty} \frac{x\,\mathrm{e}^{\frac{-x^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})} \quad \ldots\ldots \quad 4.5$$

$$\mu = \left[\frac{-e^{\frac{-x^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})}\right]_{\mathrm{k}}^{\infty} \quad \ldots\ldots \quad 4.6$$

$$\mu = \frac{e^{\frac{-\mathrm{k}^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})} \quad \ldots\ldots \quad 4.7$$

$$\mu = \frac{\phi(\mathrm{k})}{\Phi(\mathrm{k})} \quad \ldots\ldots \quad 4.8$$

The *variance* of the distribution (σ^2) is given by:

$$\sigma^2 = \int_{\mathrm{k}}^{\infty} (x-\mu)^2\,f(x)\,dx \quad \ldots\ldots \quad 4.9$$

$$= \int_{\mathrm{k}}^{\infty} x^2\,f(x)\,dx - \mu^2 \quad \ldots\ldots \quad 4.10$$

$$= \int_{\mathrm{k}}^{\infty} \frac{x^2\,\mathrm{e}^{\frac{-x^2}{2}}\,dx}{\sqrt{2\pi}\,\Phi(\mathrm{k})} - \mu^2 \quad \ldots\ldots \quad 4.11$$

$$= \left[\frac{-x\,\mathrm{e}^{\frac{-x^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})}\right]_{\mathrm{k}}^{\infty} + 1 - \mu^2 \quad \ldots\ldots \quad 4.12$$

$$= \frac{\mathrm{k}\,\mathrm{e}^{\frac{-\mathrm{k}^2}{2}}}{\sqrt{2\pi}\,\Phi(\mathrm{k})} + 1 - \mu^2 \quad \ldots\ldots \quad 4.13$$

$$= 1 + \mathrm{k}\left[\frac{\phi(\mathrm{k})}{\Phi(\mathrm{k})}\right] - \left[\frac{\phi(\mathrm{k})}{\Phi(\mathrm{k})}\right]^2 \quad \ldots\ldots \quad 4.14*$$

Example

Φ(k) = 1 – p
For p = 0,3 we have Φ(k) = 0,7.
And for Φ(k) = 0,7 we have k = −0,524002
and ϕ(k) = 0,347765.

The values of k and ϕ(k) were calculated with the normal and inverse normal distribution programme of Hewlett Packard (ST1-09A1 and ST1-09A2).

$$\sigma^2 = 1 + \mathrm{k}\left[\frac{\phi(\mathrm{k})}{\Phi(\mathrm{k})}\right] - \left[\frac{\phi(\mathrm{k})}{\Phi(\mathrm{k})}\right]^2$$

$$= 1 - 0,524002\left[\frac{0,347765}{0,700000}\right] - \left[\frac{0,347765}{0,700000}\right]^2$$

$$= 0,492855$$

With the aid of equation 4.14 population variances were calculated for different truncation ratios (p). These values are given in Table 1.

* The help of prof. A.J.B. Wiid in the derivation of this equation, is acknowledged with thanks.

TABLE 1
POPULATION VARIANCES FOR DIFFERENT DEGREES OF RESTRICTION OF RANGE OF THE UNDERLYING DISTRIBUTION

p	Φ(k)	k	ϕ(k)	σ_x^2
0,000	1,000			1,000000
0,001	0,999	–3,090522	0,003364	0,989582
0,010	0,990	–2,326785	0,026625	0,936700
0,020	0,980	–2,054189	0,048374	0,896166
0,030	0,970	–1,881213	0,067988	0,863232
0,040	0,960	–1,751077	0,086115	0,834876
0,050	0,950	–1,645211	0,103075	0,809722
0,060	0,940	–1,555097	0,119063	0,786984
0,070	0,930	–1,476078	0,134211	0,766157
0,080	0,920	–1,405322	0,148614	0,746894
0,090	0,910	–1,340969	0,162344	0,728945
0,100	0,900	–1,281729	0,175458	0,712116
0,125	0,875	–1,150436	0,205833	0,674037
0,150	0,850	–1,036431	0,233159	0,640459
0,175	0,825	–0,934503	0,257796	0,610343
0,200	0,800	–0,841457	0,280001	0,582988
0,225	0,775	–0,755178	0,299966	0,557896
0,250	0,750	–0,674189	0,317841	0,534691
0,275	0,725	–0,597405	0,333743	0,513085
0,300	0,700	–0,524002	0,347765	0,492855
0,325	0,675	–0,453333	0,359985	0,473812
0,350	0,650	–0,384877	0,370462	0,455809
0,375	0,625	–0,318200	0,379248	0,438715
0,400	0,600	–0,252933	0,386383	0,422419
0,425	0,575	–0,188756	0,391898	0,406825
0,450	0,550	–0,125381	0,395819	0,391841
0,475	0,525	–0,062545	0,398163	0,377386
0,500	0,500	–0,000000	0,398942	0,363381
0,525	0,475	0,062545	0,398163	0,349785
0,550	0,450	0,125381	0,395819	0,336593
0,575	0,425	0,188756	0,391898	0,323762
0,600	0,400	0,252933	0,386383	0,311249
0,625	0,375	0,318200	0,379248	0,299020
0,650	0,350	0,384877	0,370462	0,287034
0,675	0,325	0,453333	0,359985	0,275253
0,700	0,300	0,524002	0,347765	0,263649
0,725	0,275	0,597405	0,333743	0,252166
0,750	0,250	0,674189	0,317841	0,240773
0,775	0,225	0,755178	0,299966	0,229415
0,800	0,200	0,841457	0,280001	0,218030
0,825	0,175	0,934503	0,257796	0,206553
0,850	0,150	1,036431	0,233159	0,194883
0,875	0,125	1,150436	0,205833	0,182879
0,900	0,100	1,281729	0,175458	0,170345
0,925	0,075	1,439800	0,141501	0,156880
0,950	0,050	1,645211	0,103075	0,141820
0,960	0,040	1,751077	0,086115	0,134979
0,970	0,030	1,881213	0,067988	0,127366
0,980	0,020	2,054189	0,048374	0,118357
0,990	0,010	2,326785	0,026625	0,106159
0,995	0,005	2,576236	0,014445	0,096425
0,996	0,004	2,652463	0,011835	0,093773
0,997	0,003	2,748155	0,009140	0,090534
0,998	0,002	2,878506	0,006334	0,086340
0,999	0,001	3,090522	0,003364	0,080020

Following this, population correlations were computed for different truncation ratios.

Example

$$\rho_{xy} = \frac{\sigma_x \, \rho_{XY}}{\sqrt{1-\rho_{XY}^2(1-\sigma_x^2)}} \qquad \ldots\ldots \quad 3.13$$

From the previous example $\sigma_x^2 = 0{,}492855$. Let us consider a population correlation of 0,8 with a truncation ratio of p = 0,3, then

$$\rho_{xy} = \frac{0,702036 \times 0,8}{\sqrt{1-0,64\;(1-0,492855)}}$$

$$= \frac{0,561629}{\sqrt{1-0,64 \times 0,507145}}$$

$$= \frac{0,561629}{\sqrt{0,675427}}$$

$$= 0,683377$$

With the aid of equation 3.13 population correlations were calculated for different truncation ratios (p). These values are given in Table 2.

An interesting illustration is given in Figure 1:

The bivariate normal distribution was prepared using a method designed by Silverstein (1971). It represents a one standard deviation correlation surface of 0,6.

The standardised point of truncation on the predictor (X) is k = 1,000000, and represents a truncation ratio of p = 0,841345. In this case ϕ(k) = 0,241971. The truncated population standard deviation (σ_x) is 0,446194. And the restricted standard deviation (σ_y) of the criterion is 0,843607. The truncated population correlation is equal to 0,317347.

According to Table 2 a population correlation of 0,60 will shrink to 0,314, given a truncation ratio of 0,85. Through linear interpolation a very accurate answer can be obtained.

Figure 2 gives the residual variance of Y, conditional on X. From equation (3.2) we have

$$\sigma_y \sqrt{1-\rho_{xy}^2} = \sigma_Y \sqrt{1-\rho_{XY}^2}.$$

Substituting for the relevant values in the equation, we have:

Left hand side:

$$0,843607\sqrt{1-(0,317347)^2}$$

$$0,843607\sqrt{1-(0,100709)}$$

$$= 0,843607 \times 0,948309$$

$$= 0,800000$$

Right hand side:

$$\sqrt{1-(0,6)^2}$$

$$= 0,800000$$

Thus the standard errors of estimate of the selected and unselected groups are equal.

To obtain an overview of the effect of restriction of range on the population correlations, the truncated population correlations (ρ_p) have been plotted against the truncation ratios (p), in Figure 4.

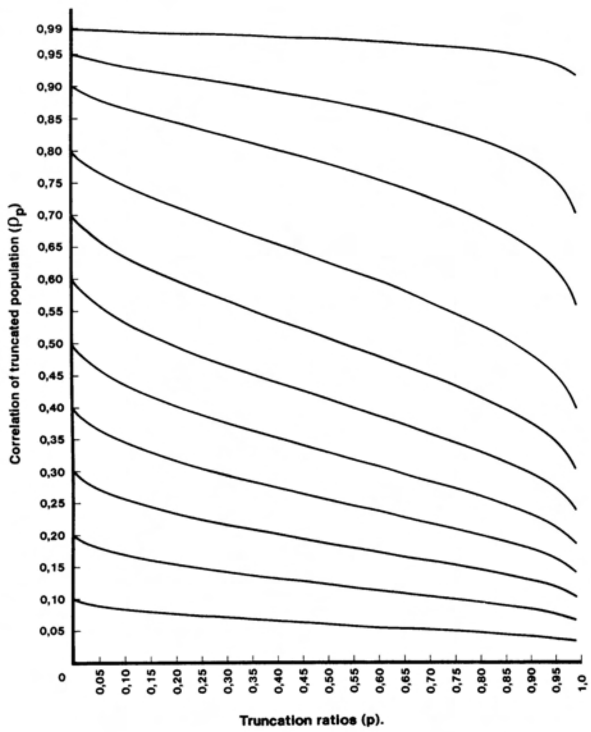

Figure 4: Correlation of truncated population (ρ_p) for various truncation ratios (p).

From an inspection of Figure 4 it should be clear that *very high* and *very low* correlations are only *slightly* influenced by restriction of range. By contrast correlations in the middle range (0,4 to 0,8) are very strongly influenced by restriction of range. It would therefore appear necessary to correct validity coefficients for explicit selection on the predictor variable. The statistical tables contained in this paper can readily be used to obtain reasonable estimates of the validities of selection tests, given that the selection ratios are known.

TABLE 2
POPULATION CORRELATIONS FOR DIFFERENT DEGREES OF RESTRICTION OF RANGE OF THE PREDICTOR VARIABLE

p	$\rho = 0,10$	$\rho = 0,15$	$\rho = 0,20$	$\rho = 0,25$	$\rho = 0,30$	$\rho = 0,35$	$\rho = 0,40$	$\rho = 0,45$	$\rho = 0,50$	$\rho = 0,55$	p
0,000	0,100	0,150	0,200	0,250	0,300	0,350	0,400	0,450	0,500	0,550	0,000
0,001	0,099	0,149	0,199	0,249	0,299	0,348	0,398	0,448	0,498	0,548	0,001
0,010	0,097	0,145	0,194	0,242	0,291	0,340	0,389	0,438	0,488	0,537	0,010
0,020	0,095	0,142	0,190	0,237	0,285	0,333	0,382	0,431	0,480	0,529	0,020
0,030	0,093	0,140	0,186	0,233	0,280	0,328	0,376	0,424	0,473	0,522	0,030
0,040	0,091	0,137	0,183	0,230	0,276	0,323	0,370	0,418	0,467	0,516	0,040
0,050	0,090	0,135	0,181	0,226	0,272	0,319	0,366	0,413	0,461	0,510	0,050
0,060	0,089	0,133	0,178	0,223	0,269	0,315	0,361	0,408	0,456	0,504	0,060
0,070	0,088	0,132	0,176	0,220	0 265	0,311	0,357	0,404	0,451	0,499	0,070
0,080	0,087	0,130	0,174	0,218	0,262	0,307	0,353	0,399	0,446	0,495	0,080
0,090	0,085	0,128	0,172	0,215	0,259	0,304	0,349	0,395	0,442	0,490	0,090
0,100	0,085	0,127	0,170	0,213	0,257	0,301	0,346	0,391	0,438	0,486	0,100
0,125	0,082	0,124	0,165	0,207	0,250	0,293	0,337	0,382	0,428	0,476	0,125
0,150	0,080	0,121	0,161	0,202	0,244	0,286	0,330	0,374	0,419	0,466	0,150
0,175	0,078	0,118	0,157	0,198	0,239	0,280	0,323	0,366	0,411	0,457	0,175
0,200	0,077	0,115	0,154	0,193	0,233	0,274	0,316	0,359	0,403	0,449	0,200
0,225	0,075	0,113	0,151	0,189	0,229	0,269	0,310	0,352	0,396	0,441	0,225
0,250	0,073	0,110	0,148	0,186	0,224	0,264	0,304	0,346	0,389	0,434	0,250
0,275	0,072	0,108	0,145	0,182	0,220	0,259	0,298	0,340	0,382	0,427	0,275
0,300	0,070	0,106	0,142	0,178	0,216	0,254	0,293	0,334	0,376	0,420	0,300
0,325	0,069	0,104	0,139	0,175	0,212	0,249	0,288	0,328	0,369	0,413	0,325
0,350	0,068	0,102	0,137	0,172	0,208	0,245	0,283	0,322	0,363	0,406	0,350
0,375	0,066	0,100	0,134	0,169	0,204	0,240	0,278	0,317	0,357	0,400	0,375
0,400	0,065	0,098	0,132	0,165	0,200	0,236	0,273	0,311	0,351	0,393	0,400
0,425	0,064	0,096	0,129	0,162	0,197	0,232	0,268	0,306	0,346	0,387	0,425
0,450	0,063	0,095	0,127	0,160	0,193	0,228	0,264	0,301	0,340	0,381	0,450
0,475	0,062	0,093	0,124	0,157	0,190	0,224	0,259	0,296	0,334	0,375	0,475
0,500	0,060	0,091	0,122	0,154	0,186	0,220	0,254	0,291	0,329	0,369	0,500
0,525	0,059	0,089	0,120	0,151	0,183	0,216	0,250	0,286	0,323	0,363	0,525
0,550	0,058	0,088	0,118	0,148	0,179	0,212	0,245	0,281	0,318	0,357	0,550
0,575	0,057	0,086	0,115	0,145	0,176	0,208	0,241	0,276	0,312	0,351	0,575
0,600	0,056	0,084	0,113	0,143	0,173	0,204	0,237	0,271	0,307	0,345	0,600
0,625	0,055	0,083	0,111	0,140	0,169	0,200	0,232	0,266	0,301	0,339	0,625
0,650	0,054	0,081	0,109	0,137	0,166	0,196	0,228	0,261	0,296	0,333	0,650
0,675	0,053	0,079	0,106	0,134	0,163	0,192	0,223	0,256	0,290	0,327	0,675
0,700	0,052	0,078	0,104	0,131	0,159	0,188	0,219	0,250	0,284	0,320	0,700
0,725	0,050	0,076	0,102	0,129	0,156	0,184	0,214	0,245	0,278	0,314	0,725
0,750	0,049	0,074	0,100	0,126	0,153	0,180	0,209	0,240	0,273	0,307	0,750
0,775	0,048	0,072	0,097	0,123	0,149	0,176	0,205	0,235	0,267	0,301	0,775
0,800	0,047	0,071	0,095	0,120	0,145	0,172	0,200	0,229	0,260	0,294	0,800
0,825	0,046	0,069	0,092	0,117	0,141	0,167	0,195	0,223	0,254	0,287	0,825
0,850	0,044	0,067	0,090	0,113	0,138	0,163	0,189	0,217	0,247	0,279	0,850
0,875	0,043	0,065	0,087	0,110	0,133	0,158	0,183	0,211	0,240	0,271	0,875
0,900	0,041	0,062	0,084	0,106	0,129	0,152	0,177	0,204	0,232	0,262	0,900
0,925	0,040	0,060	0,081	0,102	0,124	0,146	0,170	0,196	0,223	0,252	0,925
0,950	0,038	0,057	0,077	0,097	0,118	0,139	0,162	0,186	0,212	0,241	0,950
0,960	0,037	0,056	0,075	0,094	0,115	0,136	0,158	0,182	0,207	0,235	0,960
0,970	0,036	0,054	0,073	0,092	0,112	0,132	0,154	0,177	0,202	0,229	0,970
0,980	0,035	0,052	0,070	0,088	0,108	0,127	0,148	0,171	0,195	0,221	0,980
0,990	0,033	0,049	0,066	0,084	0,102	0,121	0,141	0,162	0,185	0,210	0,990
0,995	0,031	0,047	0,063	0,080	0,097	0,115	0,134	0,155	0,176	0,200	0,995
0,996	0,031	0,046	0,062	0,079	0,096	0,114	0,132	0,153	0,174	0,198	0,996
0,997	0,030	0,046	0,061	0,077	0,094	0,112	0,130	0,150	0,171	0,194	0,997
0,998	0,030	0,045	0,060	0,076	0,092	0,109	0,127	0,146	0,167	0,190	0,998
0,999	0,028	0,043	0,058	0,073	0,089	0,105	0,123	0,141	0,161	0,183	0,999

TABLE 2
POPULATION CORRELATIONS FOR DIFFERENT DEGREES OF RESTRICTION OF RANGE OF THE PREDICTOR VARIABLE
(continued)

p	ρ = 0,60	ρ = 0,65	ρ = 0,70	ρ = 0,75	ρ = 0,80	ρ = 0,85	ρ = 0,90	ρ = 0,95	ρ = 0,99	p
0,000	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,990	0,000
0,001	0,598	0,648	0,698	0,748	0,798	0,849	0,899	0,950	0,990	0,001
0,010	0,587	0,638	0,688	0,739	0,790	0,842	0,894	0,947	0,989	0,010
0,020	0,579	0,629	0,680	0,732	0,784	0,837	0,890	0,945	0,989	0,020
0,030	0,572	0,622	0,673	0,725	0,778	0,832	0,887	0,943	0,988	0,030
0,040	0,565	0,616	0,667	0,720	0,773	0,828	0,884	0,941	0,988	0,040
0,050	0,559	0,610	0,661	0,714	0,768	0,824	0,881	0,939	0,988	0,050
0,060	0,554	0,604	0,656	0,709	0,764	0,820	0,878	0,938	0,987	0,060
0,070	0,549	0,599	0,651	0,704	0,759	0,816	0,875	0,936	0,987	0,070
0,080	0,544	0,594	0,646	0,700	0,755	0,813	0,872	0,935	0,987	0,080
0,090	0,539	0,590	0,642	0,696	0,751	0,809	0,870	0,933	0,986	0,090
0,100	0,535	0,585	0,637	0,691	0,747	0,806	0,867	0,932	0,986	0,100
0,125	0,524	0,575	0,627	0,681	0,738	0,798	0,861	0,928	0,985	0,125
0,150	0,515	0,565	0,617	0,672	0,730	0,791	0,856	0,925	0,985	0,150
0,175	0,506	0,556	0,608	0,663	0,721	0,783	0,850	0,922	0,984	0,175
0,200	0,497	0,547	0,599	0,655	0,713	0,776	0,844	0,919	0,983	0,200
0,225	0,489	0,538	0,591	0,646	0,706	0,770	0,839	0,915	0,982	0,225
0,250	0,481	0,530	0,583	0,638	0,698	0,763	0,834	0,912	0,982	0,250
0,275	0,473	0,522	0,575	0,630	0,691	0,756	0,828	0,909	0,981	0,275
0,300	0,466	0,515	0,567	0,623	0,683	0,750	0,823	0,906	0,980	0,300
0,325	0,459	0,507	0,559	0,615	0,676	0,743	0,818	0,902	0,979	0,325
0,350	0,452	0,500	0,552	0,608	0,669	0,737	0,813	0,899	0,978	0,350
0,375	0,445	0,493	0,545	0,601	0,662	0,730	0,807	0,896	0,978	0,375
0,400	0,438	0,486	0,537	0,593	0,655	0,724	0,802	0,892	0,977	0,400
0,425	0,432	0,479	0,530	0,586	0,648	0,717	0,796	0,889	0,976	0,425
0,450	0,425	0,472	0,523	0,579	0,641	0,711	0,791	0,885	0,975	0,450
0,475	0,418	0,465	0,516	0,572	0,634	0,704	0,785	0,882	0,974	0,475
0,500	0,412	0,458	0,509	0,564	0,626	0,697	0,780	0,878	0,973	0,500
0,525	0,405	0,451	0,502	0,557	0,619	0,690	0,774	0,874	0,972	0,525
0,550	0,399	0,445	0,494	0,550	0,612	0,683	0,768	0,870	0,971	0,550
0,575	0,393	0,438	0,487	0,542	0,604	0,676	0,761	0,866	0,970	0,575
0,600	0,386	0,431	0,480	0,535	0,597	0,669	0,755	0,862	0,969	0,600
0,625	0,379	0,424	0,472	0,527	0,589	0,662	0,749	0,857	0,968	0,625
0,650	0,373	0,417	0,465	0,519	0,581	0,654	0,742	0,852	0,966	0,650
0,675	0,366	0,409	0,457	0,511	0,573	0,646	0,735	0,847	0,965	0,675
0,700	0,359	0,402	0,450	0,503	0,565	0,638	0,727	0,842	0,964	0,700
0,725	0,352	0,395	0,442	0,495	0,556	0,630	0,720	0,837	0,962	0,725
0,750	0,345	0,387	0,433	0,486	0,547	0,621	0,712	0,831	0,960	0,750
0,775	0,338	0,379	0,425	0,477	0,538	0,612	0,703	0,825	0,958	0,775
0,800	0,331	0,371	0,416	0,468	0,529	0,602	0,694	0,818	0,956	0,800
0,825	0,323	0,362	0,407	0,458	0,518	0,591	0,684	0,810	0,954	0,825
0,850	0,314	0,353	0,397	0,448	0,507	0,580	0,674	0,802	0,952	0,850
0,875	0,305	0,344	0,387	0,436	0,495	0,568	0,662	0,793	0,949	0,875
0,900	0,296	0,333	0,375	0,424	0,482	0,554	0,649	0,782	0,945	0,900
0,925	0,285	0,321	0,362	0,410	0,467	0,539	0,633	0,770	0,941	0,925
0,950	0,272	0,307	0,346	0,393	0,449	0,519	0,614	0,753	0,935	0,950
0,960	0,266	0,300	0,339	0,385	0,440	0,510	0,604	0,745	0,932	0,960
0,970	0,259	0,292	0,330	0,375	0,430	0,499	0,593	0,736	0,929	0,970
0,980	0,250	0,282	0,320	0,363	0,417	0,485	0,579	0,723	0,924	0,980
0,990	0,237	0,268	0,304	0,347	0,398	0,465	0,558	0,704	0,916	0,990
0,995	0,227	0,257	0,291	0,332	0,383	0,448	0,540	0,687	0,909	0,995
0,996	0,224	0,253	0,287	0,328	0,378	0,443	0,534	0,682	0,907	0,996
0,997	0,220	0,249	0,283	0,323	0,372	0,437	0,528	0,675	0,904	0,997
0,998	0,215	0,244	0,277	0,316	0,365	0,428	0,519	0,666	0,900	0,998
0,999	0,208	0,235	0,267	0,305	0,353	0,415	0,504	0,652	0,893	0,999

REFERENCES

Alexander, R.A. (1988). Group homogeneity, range restriction, and range enhancement effects on correlations. *Personnel Psychology*, 41, 773-777.

Boone, J.O. & Lewis, M.A. (1980). Demonstration of possible effects of recruitment and selection procedures on correcting the validity coefficient for restriction in range. *Psychological Reports*, 46, 927-930.

Cleary, T.A. (1968). Test bias: prediction of grades of Negro and White students in integrated colleges. Journal of *Educational Measurement*, 5(2), 115-124.

Clemans, W.V. (1956). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monograph Number 14.*

Gross, A.L. (1982). Relaxing the assumptions underlying corrections for restriction of range. *Educational and Psychological Measurement,* 42, 795-801.

Gross, A.L. & Kagen, E. (1983). Not correcting for restriction of range can be advantageous. *Educational and Psychological Measurement,* 41, 389-396.

Gulliksen, H. (1950). *Theory of Mental Tests.* New York: John Wiley & Sons, Inc.

House, J.D. (1983). Effects of restriction of range on predictive validity for the Graduate Record Examination. *Psychological Reports*, 53, 710.

Howell, D.C. (1992). *Statistical methods for psychology,* Boston: PWS-Kent Publishing Company.

Lee, R., Miller, K.J. & Graham, W.K. (1982). Corrections for restriction of range and attenuation in criterion-related validation studies. *Journal of Applied Psychology,* 67, 637-639.

Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores.* Reading, Massachusetts: Addison-Wesley Publishing Company.

Magnusson, D. (1967). *Test theory.* Reading, Massachusetts: Addison-Wesley Publishing Company.

Mendoza, J.L. & Mumford, M. (1987). Corrections for attenuation and range restriction on the predictor. *Journal of Educational Statistics,* 12, 282-293.

Pearson, K. (1903). Mathematical contributions to the theory of evolution. X1: On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society, London.* Series A, 200, 1-66.

Rydberg, S. (1962). Methods of correcting correlations for indirect restriction of range with non-interval data. *Psychometrika*, 27, 49-58.

Schepers, J.M. (1992). *Toetskonstruksie: Teorie en praktyk.* Johannesburg: RAU-Drukpers.

Silverstein, A.B. (1971). Graphing the bivariate normal distribution. *Perceptual and Motor Skills,* 33, 241-242.

Thorndike, R.L. (1949). *Personnel selection.* New York: John Wiley & Sons, Inc.