*Behavior Research Methods, Instruments, & Computers*
*1997, 29 (3), 464-467*

# VALCOR: A program for estimating standard error, confidence intervals, and probability of corrected validity

JESÚS F. SALGADO
*University of Santiago de Compostela, Santiago de Compostela, Spain*

*VALCOR is a Turbo-Basic program that corrects the observed (uncorrected) validity coefficients for criterion and predictor unreliability and range restriction in the predictor. Furthermore, using the formulas for the standard error of functions of correlations derived by Bobko and Rieck (1980), the program provides an estimation of the standard error, the confidence intervals, and the probability of the corrected validity coefficients. In this way, the probability and the boundaries of the corrected validity coefficients may be reported together with the probability of the uncorrected validity coefficients. The results are presented on the computer screen and may be saved in an external file.*

The validity of psychological measurement instruments (e.g., tests, questionnaires, interviews, etc.) is frequently estimated by using correlation coefficients. The estimated validity may be affected by a number of error sources, the most relevant being sampling error, measurement error, and range restriction (Hunter & Schmidt, 1990). The sampling error affects the correlation coefficient by reducing its efficiency in detecting significant effects. A simple way to solve this problem is to increase the sample size, thereby increasing the power of the correlation coefficient to detect the significant effect. However, the other two problems, the measurement error (unreliability) in predictor and criterion and the range restriction in criterion have different effects on validity.

The unreliability in the criterion and the predictor systematically affect validity by underestimating the true validity. The same is true for the range restriction in the criterion. Thus, the effect of these two types of errors is an attenuation of the "real" validity of the instrument. These errors may have dramatic consequences for the psychological measures. In this respect, the Society for Industrial and Organizational Psychology's (SIOP, 1987) Principles suggest that appropriate adjustments should be made when possible because the adjusted coefficient is the best point estimate of the population validity. The Principles also maintain that both unadjusted and adjusted coefficients should be reported.

I would like to thank Philip Bobko for his helpful comments on some formulas used in the program, and Roger Bakeman, Tom Nygren, and one anonymous reviewer for their suggestions on a previous version of this article. Correspondence should be addressed to J. F. Salgado, Dpto. de Psicología Social y Básica, Universidad de Santiago de Compostela, 15706 Santiago de Compostela, Spain (e-mail: psjesal@usc.es).

The adjusted coefficients for measurement error and range restriction can be easily obtained by applying "classic" psychometric formulas. Corrected validity for criterion unreliability and for predictor unreliability, or both, may be estimated by using Formulas 1, 2, and 3, for example, provided by Guilford and Fruchter (1979):

$$r_{cy} = r/(r_{yy})^{1/2}, \quad (1)$$

$$r_{cx} = r/(r_{xx})^{1/2}, \quad (2)$$

$$r_{c} = r/[(r_{yy})(r_{xx})]^{1/2}, \quad (3)$$

where $r$ = observed validity, $r_{cy}$ = validity corrected for criterion unreliability, $r_{cx}$ = validity corrected for predictor unreliability, $r_c$ = validity corrected for criterion and predictor unreliability, $r_{yy}$ = criterion reliability, and $r_{xx}$ = predictor reliability.

On the other hand, the correction for range restriction in the predictor may be estimated by using Formula 4 as provided, for example, by Ghiselli, Campbell, and Zedeck (1981):

$$r_r = r * U/(1 - r^2 + r^2 * U^2)^{1/2}. \quad (4)$$

where $r_r$ = validity corrected for range restriction in the predictor, $r$ = observed validity, and $U$ = the ratio of the standard deviation (*SD*) of the population and the *SD* of the sample ($U = SD_{pop}/SD_{sam}$).

Good estimates of the true validity may be obtained by using these four formulas. However, one problem is still present. Frequently, researchers apply significance tests to validity coefficients using the standard error (*SE*) of the correlation coefficient. Nevertheless, corrected validity coefficients are functions of correlation coefficients, and the *SE*s of these estimators are higher in magnitude than the *SE*s of the initial correlations (Bobko & Rieck, 1980). Therefore, as SIOP's Principles indicate, the usual statistical significance tests do not apply to adjusted coefficients. For this reason, as Bobko (1995) suggests, it is customary, in the literature, to test the observed validity for statistical significance and to report the corrected validity. However, there is no reason to report an untested coefficient; the only problem is to use the appropriate *SE* in the statistical test. To conduct significance tests with corrected validity coefficients, Bobko and Rieck (1980) derived the *SE* of the correlation corrected for unreliability and for range restriction.

In the case of the *SE* of the correlation corrected for measurement error in the criterion ($r_c$), there are three possible formulas, depending on the way in which the criterion or predictor reliability is obtained. In the most general situation, when the criterion or predictor reliability is an "a priori" value based on theoretical assumptions or previously accepted knowledge, the formulas for the *SE* of $r_{cy}$ and $r_{cx}$ are

$$SE(r_{cy}) = [(1 - r^2)^2/(n * r_{yy})]^{1/2}, \quad (5)$$

$$SE(r_{cx}) = [(1 - r^2)^2/(n * r_{xx})]^{1/2}, \quad (6)$$

where $SE(r_{cy})$ = $SE$ of the corrected validity for criterion unreliability, $SE(r_{cx})$ = SE of the corrected validity for predictor unreliability, $r$ = observed validity, $n$ = sample size, $r_{yy}$ = criterion reliability, and $r_{xx}$ = predictor reliability.

The other two formulas derived by Bobko and Rieck are to be used when (1) the criterion (or predictor) reliability and the validity coefficient are obtained from the same sample, and (2) when the criterion (or predictor) reliability is obtained from an independent sample. These two formulas add to Formulas 5 and 6 the variance corresponding to the sampling effect, but the dominating term is Formula 5 or Formula 6, depending on if validity is corrected for criterion or predictor unreliability.

A question regarding the use of these last formulas concerns the appropriate reliability coefficient to be used. The formulas may use all reliability coefficients, but the researcher must be conscious of what the appropriate coefficient is for each case. For example, in the area of personnel selection, when test, interviews, or other procedures are validated using performance ratings as the criterion, the appropriate reliability is the interrater coefficient; coefficients of internal consistency (e.g., Cronbach's alpha) would be incorrect in this case (Hunter & Schmidt, 1990).

The estimation of the *SE* of the validity corrected for range restriction is relatively easy to estimate, but it presents two cases. In Case A, the range restriction ratio ($U$) is known because the standard deviation (*SD*) of the population and the *SD* of the sample are known. In Case B, $U$ must be estimated from the selection ratio, defined as the number of applicants for each position. The selection ratio may be also defined as the ratio between the number of members in the restricted group and the number of members in the large group. Bobko and Reick (1980) derived the following formula to estimate the *SE* of $r_r$:

$$SE(r_r) = \{U/[1 + r^2(U^2 - 1)]^{3/2}\} * (SE_r), \quad (7)$$

where $SE(r_r)$ = $SE$ of $r_r$, $U = SD_{pop}/SD_{samp}$ ratio, $r$ = observed validity, and $SE_r$ = $SE$ of the observed validity = $1/(n)^{1/2}$.

In Case B, when $U$ is unknown but the selection ratio is known, it is possible to apply Formula 8 to estimate $U$ (Schmidt, Hunter, & Urry, 1976):

$$U = \{1/[1-(z/p) * y]-(y/p)^2\}^{1/2}, \quad (8)$$

where $z$ = "$z$" score corresponding to the selection ratio, $p$ = selection ratio, and $y$ = ordinate corresponding to $z$.

If this $U$ value is used in Formula 7, it is possible to obtain the *SE* of the validity corrected for range restriction.

Therefore, if the *SE* of the corrected validity is known, the testing of the hypothesis may be conducted, and the probability associated with the corrected validity may be reported. To estimate this probability, it is necessary to transform the corrected validity in a Student's $t$ and, after this, "$t$" in a $z$ score using the following formulas:

$$t = r\ \text{corrected}/SE_A \quad (9)$$

(Bobko, 1995),

$$z = 1/2[\log(1 + p) - \log(1 - p)] \quad (10)$$

(Downie & Heath, 1970), where $t$ = value of the "$t$" with $n - 2$ degrees of freedom, $z$ = $z$ score corresponding to the "$t$" value, $SE_A$ = Appropriate $SE$ ($SE_{rc}$ or $SE_{rr}$), and $p$ = corrected validity.

Although the hypothesis testing can be conducted using these formulas, a number of authors suggest that significance tests do not appropriately control the Type II error, and they suggest the use of confidence intervals around the effect size as an alternative (e.g., Cohen, 1994; Hunter & Schmidt, 1990). To develop the confidence intervals, the *SE*s are also necessary, and they should be reported with the corrected coefficients.

From an applied point of view, the only limitation of these procedures is the application of formulas, because some of them are very difficult to calculate without a computer. For example, the computation of the *SE* of the corrected validity for criterion unreliability (Formulas 5, 6, and variations) or Formula 7 required considerable computational effort and work. The same argument may be applied to Formulas 8, 9, and 10 and the estimation of the probability of corrected validity. Moreover, these formulas are not commonly seen in the literature. This may be another explanation for the nontesting of the corrected validity. Unfortunately, there are no computer programs for this currently available for researchers and practitioners.

This paper describes VALCOR, a program written in Turbo-Basic that computes validity corrected for measurement error in the criterion and for range restriction in the predictor (Cases A and B) and the *SE*s, probabilities, and confidence intervals of the corrected and uncorrected validities. To illustrate the computations made by the program, two outputs of the application of VALCOR to the example presented by Bobko (1995, pp. 81 and 110) are presented.

**Program Description**

The program VALCOR is composed of three subprograms that mainly perform the estimation of (1) validity coefficient corrected for criterion unreliability, (2) validity coefficient corrected for predictor unreliability, (3) validity corrected for both criterion and predictor unreliability, (4) validity coefficient corrected for range restriction in predictor, (5) *SE* of the observed and corrected validity coefficients, (6) confidence intervals for uncorrected and corrected validity coefficients (the user chooses the desired interval), and (7) the probability associated with the uncorrected and corrected validity coefficients. The estimation of these coefficients is made using the formulas presented in the introduction to this article.

For the first subprogram (to be used when the correction of validity is for measurement error in the criterion), the user must indicate the uncorrected validity, the sample size, the criterion reliability, the predictor reliability, how the reliabilities were obtained (e.g., in the same sample, in another sample, or a theoretical value), and the confidence interval desired.

For the second subprogram (to be used when the correction of validity is for range restriction in the predictor), the user must indicate the uncorrected validity, the sample size, the selection ratio, and the confidence interval desired.

For the third subprogram (to be used when the correction of validity is for range restriction in the predictor), the user must indicate the uncorrected validity, the sample size, the range restriction ratio ($U$), and the confidence interval desired.

All the estimations calculated for the program are presented on the screen, but the user has the option of filing the output in an external file. The program checks the values entered by the user, and it runs only if the values are in the appropriate range (e.g., 0 to 1 for reliability, −1 to +1 for validity, etc.). Also, the program contains a Help utility which explains both the main concepts used in the formulas and how to run the program.

**Table 1**
**VALCOR Output for Example 1 (Bobko, 1995, p. 81)**

| | | |
|---|---|---|
| VALCOR – A PROGRAM FOR VALIDITY CORRECTION | | |
| RESULTS OF VALIDITY CORRECTION FOR MEASUREMENT ERRORS | | |
| PROBLEM TITLE = Output of EX1. | | |
| DATE = 7/6/96 | | |
| OBSERVED VALIDITY (UNCORRECTED) R | = | −0.170 |
| SAMPLE SIZE N | = | 254 |
| CRITERION RELIABILITY RYY | = | 0.87 |
| PREDICTOR RELIABILITY RXX | = | 1.00 |
| RCY (VALIDITY CORRECTED FOR RYY) | = | −0.182 |
| RCX (VALIDITY CORRECTED FOR RXX) | = | −0.170 |
| VALIDITY CORRECTED FOR RYY AND RXX | = | −0.182 |
| STANDARD ERROR FOR R | = | 0.0609 |
| STANDARD ERROR FOR RCY | = | 0.0701 |
| STANDARD ERROR FOR RCX | = | 0.0609 |
| CONFIDENCE INTERVAL (CI) | = | 90% |
| CONFIDENCE INTERVAL FOR R | | |
| LOWER BOUNDARY | = | −0.270 |
| HIGHER BOUNDARY | = | −0.070 |
| CONFIDENCE INTERVAL FOR RCY | | |
| LOWER BOUNDARY | = | −0.298 |
| HIGHER BOUNDARY | = | −0.067 |
| CONFIDENCE INTERVAL FOR RCX | | |
| LOWER BOUNDARY | = | −0.270 |
| HIGHER BOUNDARY | = | −0.070 |
| PROBABILITY OF R (ONE TAILED TEST) | = | 0.0033 |
| CONFIDENCE LEVEL | = | 0.9967 |
| PROBABILITY OF R (TWO TAILED TEST) | = | 0.0066 |
| CONFIDENCE LEVEL | = | 0.9934 |
| PROBABILITY OF RCY (ONE TAILED TEST) | = | 0.0050 |
| CONFIDENCE LEVEL | = | 0.9950 |
| PROBABILITY OF RCY (TWO TAILED TEST) | = | 0.0099 |
| CONFIDENCE LEVEL | = | 0.9901 |
| PROBABILITY OF RCX (ONE TAILED TEST) | = | 0.0080 |
| CONFIDENCE LEVEL | = | 0.9920 |
| PROBABILITY OF RCX (TWO TAILED TEST) | = | 0.0161 |
| CONFIDENCE LEVEL | = | 0.9839 |

Note—If the confidence interval boundaries are higher than 1.00 or lower than −1.00, it is highly probable that the distribution is nonnormal bivariate.

**Table 2**
**VALCOR Output for Example 2 (Bobko, 1995, p. 104)**

| | | |
|---|---|---|
| VALCOR–A PROGRAM FOR VALIDITY CORRECTION | | |
| RESULTS OF VALIDITY CORRECTION FOR RANGE RESTRICTION | | |
| PROBLEM TITLE = Ouput of EX2 | | |
| DATE = 06-07-1996 | | |
| OBSERVED VALIDITY (UNCORRECTED) R = 0.110 | | |
| SAMPLE SIZE N = 44 | | |
| SDpop/SDres RATIO U = 5.000 | | |
| STANDARD ERROR (SE) OF R = 0.1489 | | |
| VALIDITY CORRECTED FOR RANGE RESTRICTION RR = 0.4842 | | |
| STANDARD ERROR (SE) OF RR = 0.5080 | | |
| STANDARD ERROR OF RR IN SDr UNITS = 3.41 | | |
| CONFIDENCE INTERVAL (CI) | = | 90 % |
| Z VALUE FOR CI | = | 1.644 |
| CONFIDENCE INTERVAL FOR R | | |
| LOWER BOUNDARY | = | −0.135 |
| HIGHER BOUNDARY | = | 0.355 |
| CONFIDENCE INTERVAL FOR RR | | |
| LOWER BOUNDARY | = | −0.351 |
| HIGHER BOUNDARY | = | 1.320 |
| PROBABILITY OF R (ONE TAILED TEST) | = | 0.2386 |
| CONFIDENCE LEVEL | = | 0.7614 |
| PROBABILITY OF R (TWO TAILED TEST) | = | 0.4772 |
| CONFIDENCE LEVEL | = | 0.5228 |
| PROBABILITY OF RR (ONE TAILED TEST) | = | 0.1730 |
| CONFIDENCE LEVEL | = | 0.8270 |
| PROBABILITY OF RR (TWO TAILED TEST) | = | 0.3460 |
| CONFIDENCE LEVEL | = | 0.6540 |

Note—If the confidence interval boundaries are higher than 1.00 or lower than −1.00, it is highly probable that the distribution is nonnormal bivariate.

## Examples

VALCOR was applied to two examples presented by Bobko (1995). The first example (EX1) shows the correlation between negative affect ($Y$) and prosocial behavior within groups ($X$). The validity coefficient is equal to −.17, and the criterion reliability is .87 (George, 1990, as cited by Bobko, 1995, p. 81). The second example (EX2) shows a case of range restriction in which the range restriction ratio ($U$) is 5 and the observed validity is .11 (Bobko, 1995, p. 104). The results of EX1 and EX2 are presented in Tables 1 and 2. Even though Bobko (1995, pp. 81 and 104) does not give complete results for these cases, his results are identical to the coefficients computed by the VALCOR program.

### Hardware and Software Requirements

The current VALCOR version requires an IBM-compatible PC with MS-DOS 3.3 or higher. Also, it may be used with DOS windows in Windows 3.1 or higher. I recommend installing the program on a 386 (or better) PC. Although VALCOR may run on some PCs with monochrome monitors, a color monitor is recommended.

### Availability

An EXE version of the program VALCOR is available at no charge by sending a DOS-formatted floppy disk to Jesús F. Salgado, Dpto. de Psicología Social y Básica, Universidad de Santiago de Compostela (Spain) (e-mail: psjesal@usc.es).

Considerable effort has been directed toward making VALCOR error free, but there is no warranty. Users are kindly asked to communicate with the author regarding any problem encountered with the program.

## REFERENCES


Bobko, P. (1995). *Correlation and regression: Principles and applications for industrial/organizational psychology and management.* New York: McGraw-Hill.

Bobko, P., & Rieck, A. (1980). Large sample estimators for standard errors of functions of correlation coefficients. *Applied Psychological Measurement*, **4**, 385-398.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, **49**, 997-1003.

Downie, N. M., & Heath, R. W. (1970). *Basic statistical methods* (3rd ed). New York: Harper & Row.

Ghiselli, E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W. H. Freeman.

Guilford, J. P., & Fruchter, B. (1979). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.

Hunter, J. E., & Schmidt, F L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Schmidt, F. L., Hunter, J. E., & Urry, V. W. (1976). Statistical power in criterion-related validation studies. *Journal of Applied Psychology*, **61**, 473-485.

Society for Industrial and Organizational Psychology, Inc. (1987). *Principles for the validation and use of personnel selection procedures* (3rd ed.). College Park, MD: Author.


(Manuscript received March 13, 1996;
revision accepted for publication June 13, 1996.)