

Corrections for attenuation and corrections for range restriction



One of the most pervasive methodological problems in the educational and psychological field entails determination of the techniques which are to be used in assessing the nature and strength of the relationship between various measures. Of course, the correlation coefficient has provided the field with a viable statistical tool for solving this problem. Unfortunately, in some instances the appropriateness of correlational techniques may be limited by the operation of certain statistical biases in actual data bases. Thorndike (1949) has noted that two of these biases, termed range restriction and attenuation effects, can exert a powerful diminishing influence on the magnitude of observed correlation coefficients.

Range restriction occurs when a researcher wants to estimate the correlation between two variables (x and y) in a population, but subjects are selected on x , and data for y are only available for a selected sample (Raju & Brand, 2003). This occurs for example when scores from admission tests are used to predict academic success in higher education or are compared with grades in the program they were admitted to (Gulliksen, 1950; Thorndike, 1949). Because selection is made on the basis of scores from these kinds of instruments, the range of scores is restricted in the sample. Although the correlation between test scores and academic success can be obtained for the restricted sample, the correlation for the population of applicants remains unknown. Due to the range restriction in test scores, the correlation obtained is expected to be an underestimate of the correlation in the population (Hunter & Schmidt, 1990; Henriksson & Wolming, 1998).

Attenuation effects refer to the fact that an observed correlation coefficient will tend to underestimate the true magnitude of the relationship between

<https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

two variables to the extent that these measures are not an accurate reflection of true variation, i. e., to the extent that they are unreliable. In some applied studies, the operation of these biases may be acceptable. Yet when an investigation centers on determining the true strength of the relationship between two sets of measures, the operation of these biases in the experimental data base constitutes a serious, often unavoidable, confound (Crocker & Algina, 1986; Worthen, White, Fan, & Sudweeks, 1999).

Psychometrics has long been aware of the implications of range restriction and attenuation effects with respect to the inferences drawn by researchers concerning the magnitude of relationships. Consequently, a variety of formulas have been derived which permit the researcher to correct data based estimates of the magnitude of a correlation coefficient for the operation of these influences (Guilford, 1954; Stanley, 1971). The aim of this review is to discuss the importance of correcting for range restriction and correcting for attenuation in predictive validity studies and review two methods to correction for range restriction (Thorndike's case II and ML estimates obtained from the EM algorithm) and two methods to correction for attenuation (traditional approach and latent variable modeling approach). Results from research evaluating the use of these methods will also be discussed.

Importance of corrections for range restriction and attenuation effects

As early as the beginning of the last century, Pearson (1903), in developing the Pearson product-moment correlation coefficient, noticed problems due to range restriction and attenuation and discussed possible solutions. Since <https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

then, a great number of studies have examined the biasing effect of these statistical artifacts (e. g., Alexander, 1988; Dunbar & Linn, 1991; Lawley, 1943; Linn, Harnisch, & Dunbar, 1981; Schmidt, Hunter, & Urry, 1976; Thorndike, 1949; Sackett & Yang, 2000). It is evident from literature that both range restriction and attenuation can create serious inaccuracies in empirical research, especially in the fields of employment and educational selection.

The need for correcting validity coefficients for statistical artifacts is becoming more recognized. Validity generalization research has demonstrated that artifacts like range restriction and attenuation account for large percentages of the variance in distributions of validity coefficients. Although the Society for Industrial and Organizational Psychology's (SIOP) Principles (1987) recommend correcting validity coefficients for both range restriction and criterion unreliability, researchers rarely do so. Ree et al. (1994) discussed the application of range restriction corrections in validation research. They reviewed validity articles published in *Educational and Psychological Measurement*, *Journal of Applied Psychology*, and *Personnel Psychology* between 1988 and 1992. Ree et al. (1994) concluded that only 4% of the articles dealing with validation topics applied range restriction corrections.

Researchers may be reluctant to apply corrections for range restriction and attenuation for several reasons. Seymour (1988) referred to statistical corrections as "hydraulic", implying that researchers can achieve a desired result by "pumping up" the corrections. Another reason for reluctance in applying corrections may be because the APA Standards (1974) stated that <https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

correlations should not be doubly corrected for attenuation and range restriction. The more current Standards (1985), however, endorse such corrections. A third reason for not using the corrections is that knowledge of unrestricted standard deviations is often lacking (Ree et al., 1994). Finally, researchers may be concerned that in applying corrections to correlation coefficients, they may inadvertently overcorrect.

Linn et al. (1981) stated that, “ procedures for correcting correlations for range restriction are desperately needed in highly selective situations (i. e., where selection ratios are low)” (p. 661). They continued, “ The results also clearly support the conclusion that corrections for range restriction that treat the predictor as the sole explicit selection variable are too small. Because of this undercorrection, the resulting estimates still provide a conservative indication of the predictive value of the predictor” (p. 661). Linn et al. stated that ignoring range restriction and/or attenuation corrections because they may be too large is overly cautious. They suggested the routine reporting of both observed and corrected correlations. Both observed and corrected correlations should be reported because there is no significance test for corrected correlations (Ree et al., 1994).

Based on the logic and suggestions from literature, there appear to be a number of reasons to correct for restriction of range and attenuation in predictive validity studies. These corrections could be used to adjust the observed correlations for biases, and thus yield more accurate results.

Correction Methods for Range Restriction

There are several methods for correcting correlations for range restriction.

This review is meant to examine two approaches to correction for range restriction; Thorndike's case II and ML estimates obtained from the EM algorithm. These methods will be described first, and then results from research evaluating their use will be discussed.

Thorndike's case II

Thorndike's (1949) Case II is the most commonly used range restriction correction formula in an explicit selection scenario. Explicit selection is a process, based on the predictor x , that restricts the availability of the criterion y . The criterion is only available (measured) for the selected individuals. For example, consider the seemingly straightforward case where there is direct selection on x (e. g., no one with a test score below a specified cutoff on x is selected into the organization) (Mendoza, 1993). Thorndike's Case II equation can be written as follows

$R_{xy} =$

where R_{xy} = the validity corrected for range restriction; r_{xy} = the observed validity in the restricted group; and $u_x = s_x/S_x$, where s_x and S_x are the restricted and unrestricted SDs of x , respectively. Both the restricted and unrestricted SDs of x are available at hand.

The use of this formula requires that the unrestricted, or population, variance of x be known. Although often this is known, as in the case of a predictive study where all applicants are tested and test data on all applicants are

retained, it is not uncommon to encounter the situation in which test data on applicants who were not selected are discarded and thus are not available to the researcher who later wishes to correct the sample validity coefficient for range restriction (Sackett and Yang, 2000).

Issues with Thorndike's case II method

Thorndike's Case II is by far the most widely used correction method. It is appropriate under the condition of direct range restriction (a situation where applicants are selected directly on test scores). Researchers used it and proved its appropriateness. For example, Chernyshenko and Ones (1999) and Wiberg and Sundström (2009) showed that this formula produced close estimates of correlation in a population.

Although the use of Thorndike's Case II formula is straightforward, this formula imposes some requirements. First, it requires that the unrestricted, or population, variance of x be known. Second, the formula requires that there is no additional range restriction on additional variables. If the organization also imposes an additional cutoff, such as a minimum education requirement, applying the Case II formula produces a biased result. In this example, if education level (z) and test score (x) are known for all applicants, a method for solving the problem exists (Aitken, 1934). Third, the correction formula requires two assumptions: that the x - y relationship is linear throughout the range of scores (i. e., the assumption of linearity) and that the error term is the same in the restricted sample and in the population (i. e., the assumption of homoscedasticity). Note that no normality assumption is required for the formula (Lawley, 1943).

Another issue that was found in literature with this method arises when it is applied for indirect restriction of range (a case where the applicants are selected on another variable that is correlated with the test scores) even though it has been shown to underestimate validity coefficients (Hunter & Schmidt, 2004, Ch. 5; Hunter et al., 2006; Linn et al., 1981; Schmidt, Hunter, Pearlman, & Hirsh, 1985, p. 751).

Maximum Likelihood estimates obtained from the Expectation Maximization algorithm

Using this approach, the selection mechanism is viewed as a missing data mechanism, i. e. the selection mechanism is viewed as missing, and the missing values are estimated before estimating the correlation. By viewing it as a special case of missing data, we can borrow from a rich body of statistical methods; for an overview see e. g. Little & Rubin (2002), Little (1992) or Schafer & Graham (2002). There are three general missing data situations; MCAR, MAR and MNAR. Assume X is a variable that is known for all examinees and Y is the variable of interest with missing values for some examinees. MCAR means that the data is Missing Completely At Random, i. e. the missing data distribution does not depend on the observed or missing values. In other words, the probability of missingness in data Y is unrelated to X and Y. MAR means that the data is Missing At Random, i. e. the conditional distribution of data being missing given the observed and missing values depends only on the observed values and not on the missing values. In other words, the probability of missingness in data Y is related to X, but not to Y. MNAR means that data is Missing Not At Random. In other words, the probability of missingness on Y is related to the unobserved

values of Y (Little & Rubin, 2002; Schafer & Graham, 2002). If the data is either MCAR or MAR, we can use imputation methods to replace missing data with estimates. In predictive studies, the selection mechanism that is based solely on X, the data is considered to be MAR (Mendoza, 1993). Using this approach, we can use information on some of the other variables to impute new values. Herzog & Rubin (1983) stated that by using imputation one can apply existing analysis tools to any dataset with missing observations and use the same structure and output.

There are several different techniques that use imputation to replace missing values. The most commonly applied techniques are mean imputation, hot-deck imputation, cold-deck imputation, regression imputation and multiple imputations (Madow, Olkin, & Rubin, 1983; Särndal, Swensson, & Wretman, 1992). In general, imputation may cause distortions in the distribution of a study variable or in the relationship between two or more variables. This disadvantage can be diminished when e. g. multiple regression imputation is used (Särndal et al., 1992). For example, Gustafsson & Reuterberg (2000) used regression to impute missing values in order to get a more realistic view of the relationship between grades in upper secondary schools in Sweden and the Swedish Scholastic Achievement Test. Note that regression imputation is questionable to use, because all imputed values fall directly on the regression line, the imputed data lack variability that would be present had both X and Y been collected. In other words the correlation would be 1. 0 if only computed with imputed values (Little & Rubin, 2002). Therefore literature suggest using imputed Maximum Likelihood (ML) estimates for the

missing values that are obtained using the Expectation Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977).

Maximum likelihood (ML) estimates obtained from the Expectation Maximization (EM) algorithm is imputed for the criterion variable for examinees who failed the selection test for example (Dempster et al., 1977; Little, 1992). The complete and incomplete cases were used together as the EM algorithm reestimates means, variances and covariances until the process converges. The base of EM missing values is an iterative regression imputation. The final estimated moments are the EM estimates including estimates for the correlation. For an extensive description see SPSS (2002). The idea is that the missing Y values are imputed using the following equation

where and are the estimates obtained from the final iteration of the EM algorithm. Schaffer and Graham (2002) suggested that using EM imputation is valid when examining missing data.

Issues with ML estimates obtained from the EM algorithm method

This approach is seldom used with range restriction problems, although it has been mentioned as a possibility (Mendoza, 1993). In a more recent study, Mendoza, Bard, Mumford, & Ang, (2004) concluded that the ML estimates obtained from the EM algorithm procedure produced far more accurate results. Wiberg and Sundström (2009) evaluated this approach in an empirical study and their results indicated that ML estimates obtained

from the EM algorithm seem to be a very effective method of estimating the population correlation.

Since there is not much work in literature examining the appropriateness and effectiveness of this approach, many questions need to be answered when using ML estimates obtained from the EM algorithm for correction for range restriction. Many researches need to evaluate the use of this approach in areas that are of special interest include simulations of different population correlations and different selection proportions when using the missing data approach. Regarding the EM imputation approach, one important research question is how many cases can be imputed[1]at the same time as we obtain a good estimate of the population correlation.

Correction Methods for Attenuation

In educational and psychological research, it is well known that measurement unreliability, that is, measurement error, attenuates the statistical relationship between two composites (e. g., Crocker & Algina, 1986; Worthen, White, Fan, & Sudweeks, 1999). In this review, two approaches for correcting attenuation effects caused by measurement error; traditional approach and latent variable modeling approach, will be described and results from research evaluating their use will be discussed.

Traditional approach

In classical test theory, the issue of attenuation of correlation between two composites caused by measurement unreliability is usually discussed within the context of score reliability and validity. More specifically, if there are two measured variables x and y , their correlation is estimated by the Pearson

correlation coefficient r_{xy} from a sample. Because the measured variables x and y contain random measurement error, this correlation coefficient r_{xy} is typically lower than the correlation coefficient between the true scores of the variables T_x and T_y (r_{T_x, T_y}) (Fan, 2003). When Spearman first proposed the correction for attenuation, he advocated correcting for both the predictor and the criterion variables for unreliability. His equation,

$$r_{T_x, T_y} = ,$$

is known as double correction. The double correction performed on the obtained validity coefficient reveals what the relationship would be between two variables if both were measured with perfect reliability. Because measurement error truncates, or reduces, the size of the obtained validity coefficient, the effect of the correction is to elevate the magnitude of the corrected validity coefficient above the magnitude of the obtained validity coefficient. The lower the reliability of the predictor and/or criterion variables, the greater will be the elevation of the correction. If both the test and the criterion exhibit very high reliability, the denominator of the equation will be close to unity, thus $r_{T_x, T_y} \hat{=} r_{xy}$.

The double correction formula was followed by the single correction formula as researchers began to shift the emphasis from test construction to issues of using tests to predict criteria. As the name implies, the formula involves correcting for unreliability in only one of the two variables. The formula would be either $r_{T_x, T_y} =$ (correcting for unreliability in the criterion variable only) or $r_{T_x, T_y} =$ (correcting for unreliability in the predictor variable only).

The rationale for the single correction of the criterion unreliability was best stated by Guilford (1954):

In predicting criterion measures from test scores, one should not make a complete [double] correction for attenuation. Corrections should be made in the criterion only. On the one hand it is not a fallible criterion that we should aim to predict, including all its errors; it is a 'true' criterion or the true component of the obtained criterion. On the other hand, we should not correct for errors in the test, because it is the fallible scores from which we must make predictions. We never know the true scores from which to predict. (p. 401)

Although most researchers have adopted Guilford's position on correcting only for criterion unreliability, there have been cases where correcting only for unreliability in the predictor was used. However, these occasions appear to be special cases of double correction, where either the reliability of the criterion was unknown or where the criterion was assumed to be measured with perfect reliability. The former situation was not unusual. We often know more about the reliability of tests than the reliability of criteria. The later situation is more unusual in that variables are rarely assessed with perfect reliability.

Issues with traditional approach

The correction for attenuation due to measurement error is one of the earliest applications of true-score theory (Spearman, 1904) and has been the subject of numerous debates, spurring criticisms from its very inception (e. g., Pearson, 1904). Despite this, no real consensus on correction for

attenuation has emerged in the literature, and many ambiguities regarding its application remain. One of the early criticisms is corrected validity coefficients greater than one.

Although it is theoretically impossible to have a validity coefficient in excess of 1.00, it is empirically possible to compute such a coefficient using Spearman correction formula. For example, if $r_{Tx, Ty} = .65$, $r_{Tx, Tx} = .81$, and $r_{Ty, Ty} = .49$,

$$r_{Tx, Ty} = 1.03$$

The value of 1.03 is theoretically impossible because valid variance would exceed obtained variance (error variance). Psychometricians have offered various explanations for this phenomenon. Before the year ended, Karl Pearson (1904, in his appendix) had declared that any formula that produced correlation coefficients greater than one must have been improperly derived; however, no errors were subsequently found in Spearman's formula. This led to debate over both how correction for attenuation could result in a correlation greater than one and whether a procedure that often resulted in a correlation greater than one was valid. Many explanations for correction for attenuation's supposed flaw have been suggested.

Error in estimating reliability. Many statistics used to estimate reliability are known to regularly underestimate reliability (i. e., overestimate the amount of error; Johnson, 1944; Osburn, 2000). Whereas this bias is tolerated as being in the "preferred direction" for some applications (as when a researcher wants to guarantee a minimum reliability), the result of correction for attenuation is inflated if the denominator entered into the equation is less than the accurate value (Winne & Belfry, 1982). Other researchers have
<https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

shown that some reliability estimates can overestimate reliability when transient errors are present; however, it has been argued that this effect is probably small in practice (Schmidt & Hunter, 1996, 1999).

Normal effects of sampling process. Others, including Spearman (1910), have attempted to explain corrected correlations greater than one as the normal result of sampling error. Worded more explicitly, this asserts that a corrected correlation of 1.03 should fall within the sampling distribution of corrected correlations produced by a population with a true-score correlation less than or equal to one. Despite this, it was some time before researchers first began to examine the sampling distributions of corrected correlations. However, some early studies that have examined the accuracy of correction for attenuation are of note[3].

Misunderstanding of random error. Thorndike (1907) applied multiple simulated error sets to a single set of true-score values and concluded that the equation for correction for attenuation worked reasonably well. Johnson (1944) extended this study and demonstrated that random errors would occasionally raise the level of observed correlations above the true-score correlation. In those cases, the equation to correct for attenuation corrects in the wrong direction. Johnson's conclusion that "Corrected coefficients greater than one are caused by fluctuations in observed coefficients due to errors of measurement and not by fluctuations caused by errors of sampling, as suggested by Spearman" (Johnson, 1944, p. 536). Garside (1958) referenced the various bases of error variance in the coefficients as "function fluctuations".

Latent variable modeling approach

Latent variable approach is considered when a multifactorial test is used in the admission of students to various schools. Most often a composite measure related to the total test score or subtests are used in such prediction. The use of a multiple factor latent variable model for the observed variables comprising the test can make more efficient use of the test information.

Correctly assessing the predictive validity in traditional selection studies, without latent variables, is a difficult task involving adjustments to circumvent the selective nature of the sample to be used for the validation. Latent variable modeling of the components of a test in relation to a criterion variable provides more precise predictor variables, and may include factors which have a small number of measurements. For many ability and aptitude tests it is relevant to postulate a model with both a general factor influencing all components of the test, and specific factors influencing more narrow subsets (Fan, 2003).

In confirmatory factor analysis where each latent factor has multiple indicators, measurement errors are explicitly modeled in the process. The relationship between such latent factors can be considered as free from the attenuation caused by the measurement error. For example, The GMAT exam is a standardized assessment that helps business schools assess the qualifications of applicants for advanced study in business and management. The GMAT exam measures three areas; Verbal, Quantitative Reasoning, and Analytical Writing Skills. To illustrate the point, let's look at the verbal exam. The verbal exam measures three related latent variables (Critical Reasoning <https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

(), Reading Comprehension (), Grammar and Sentence Structure ()). Each of these variables has many indicators. In such model, is considered to represent the true relationship between the three latent variables (, ,, respectively) that is not attenuated by the measurement error (to). This approach for obtaining measurement-error-free relationship between factors is well-known in the area of structural equation modeling but is rarely discussed within the context of measurement reliability and validity.

Using this approach, once the interitem correlation is obtained, the population reliability in the form of Cronbach's coefficient alpha[4] could be obtained. Cronbach's coefficient alpha takes the form

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma^2} \right)$$

where k is the number of items within a composite, $\sum \sigma_i^2$ is the sum of item variances, and σ^2 is the variance of the composite score. The variance of the composite is simply the sum of item variances () and the sum of item covariances (2).

$$\sigma^2 = \sum \sigma_i^2 + 2 \sum_{i < j} \sigma_{ij}$$

The population intervariable correlation is obtained from the two-factor model in the Figure above based on the following (Jöreskog & Sörbom, 1989):

$$\hat{\Sigma} = \hat{\Gamma} \hat{\Gamma}' \hat{\Psi}^2 + \hat{\Sigma}^e$$

where $\hat{\Sigma}$ is the population covariance matrix (correlation matrix for our standardized variables), $\hat{\Gamma}$ is the matrix of population pattern coefficients, $\hat{\Psi}^2$

is the population correlation matrix for the two factors, and $\hat{\Sigma}$ is the covariance matrix of population residuals for the items.

Issues with latent variable modeling approach

This approach for obtaining measurement-error-free correlation coefficients is well known in the area of structural modeling, but it is rarely discussed within the context of measurement reliability and validity. Fan (2003) used this approach to correct for attenuation and showed that this approach provided not only near identical and unbiased means but also near identical confidence intervals for the sampling distribution of the corrected correlation coefficients. It is pointed out, however, that the latent variable modeling approach may be less applicable in research practice due to more difficult data conditions at the item level in research practice. DeShon (1998) stated that latent variable modeling approach provides a mathematically rigorous method for correcting relationships among latent variables for measurement error in the indicators of the latent variables. However, this approach can only use the information provided to correct for attenuation in a relationship. It is not an all-powerful technique that corrects for all sources of measurement error.

Conclusion

It has long been recognized that insufficient variability in a sample will restrict the observed magnitude of a Pearson product moment coefficient. Since R. L. Thorndike's days, researchers have been correcting correlation coefficients for attenuation and/or restriction in range. The topic has received considerable attention (Bobko, 1983; Callender & Osborn, 1980; Lee, Miller, & Graham, 1982; Schmidt & Hunter, 1977) and today correlation coefficients <https://assignbuster.com/corrections-for-attenuation-and-corrections-for-range-restriction/>

are corrected for attenuation and range restriction in a variety of situations. These include test validation, selection, and validity generalization studies (meta-analysis; Hedges & Olkin, 1985), such as those conducted by Hunter, Schmidt, and Jackson (1982). For example, Pearlman, Schmidt, and Hunter (1980) corrected the mean correlation coefficient in their validity generalization study of job proficiency in clerical occupations for predictor and criterion unreliability as well as for range restriction on the predictor.

There are several methods that can be used to correct correlations for attenuation and range restriction, and some have been more frequently used than others. For correction for attenuation, the traditional method for correcting for attenuation is the best known and is easy to use. However, in more complex modeling situations it is probably easier to adopt an SEM approach to assessing relationships between variables with measurement errors 'removed' than to try to apply the traditional formula on many relationships simultaneously. Fan (2003) shows that the SEM approach (at least in the CFA context) produces equivalent results to the application of the traditional method. For correction for range restriction, the Thorndike case II method has been shown to produce close estimates of the correlation in a population (Hunter & Schmidt, 1990). Wiberg and Sundström (2009) show that ML estimates obtained from the EM algorithm approach provides a very good estimate of the correlation in the unrestricted sample as well. However, because the ML estimates obtained from the EM algorithm approach is not commonly used in range restriction studies, the usefulness and accuracy of this method should be further examined.

Using an appropriate method for correcting for attenuation and range restriction is most important when conducting predictive validity studies of instruments used, for example, for selection to higher education or employment selection. The use of inappropriate methods for statistical artifacts correction or no correction method at all could result in invalid conclusions about test quality. Thus, carefully considering methods for correcting for attenuation and range restriction in correlation studies is an important validity issue. The literature reviewed here clearly suggests that practitioners should apply attenuation and range restriction corrections whenever possible, even if the study does not focus on measurement issues (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).