Spring 5-16-2014

Adjusting Multiple Correlations for Regression Overfitting and Indirect Range Restriction

Charles Bolton III
Western Kentucky University, charles.bolton251@topper.wku.edu

ADJUSTING MULTIPLE CORRELATIONS FOR REGRESSION OVERFITTING AND INDIRECT RANGE RESTRICTION

A Capstone Experience/Thesis Project

Presented in Partial Fulfillment of the Requirements for

the Degree Bachelor of Arts with

Honors College Graduate Distinction at Western Kentucky University

By

Charles Leonard Bolton, III

*****

Western Kentucky University
2014

CE/T Committee:

Dr. Reagan D. Brown, Advisor

Dr. Anthony Paquin

Dr. Leslie Baylis

Approved by

____
Advisor
Department of Psychology

ABSTRACT

No research to date has been conducted to investigate the efficacy of and proper procedures for adjusting multiple correlations for the combined effects of regression overfitting and indirect range restriction. The present study uses Monte Carlo analyses to investigate the implementation of both of these adjustments.

Keywords: indirect range restriction, regression overfitting.

Dedicated to my amazing fiancée, Elisabeth McDermott.

ACKNOWLEDGEMENTS

I would like to thank Dr. Reagan Brown for being my advisor and mentor on this project. His hard work and dedication have made this possible. I would also like to thank Dr. Tony Paquin and Dr. Leslie Baylis for being on my advising committee for this project. Their input has been a major contribution. Finally, I want to thank my family for supporting me throughout this process - for coming to my presentation at the WKU Student Research Conference and for listening to me explain my project numerous times; thanks Elisabeth, Dad, Vince, and Mom!

VITA

January 30, 1992 .... Born-Bowling Green, Kentucky

2010 .... Bowling Green High School,
Bowling Green, Kentucky

2010 .... Accepted into
Western Kentucky University
Honors College

2010-2014 .... Regents Scholarship
Western Kentucky University

PRESENTATIONS

Bolton, C.L., Brown, R.D. (2014) Adjusting Multiple Correlations for Regression Overfitting and Indirect Range Restriction. Paper presented at 2014 Annual Student Research Conference at Western Kentucky University. Bowling Green, KY.

Bolton, C. L., Brown, R. D. (2014) Adjusting Multiple Correlations for Regression Overfitting and Indirect Range Restriction. Paper defended to Western Kentucky University's Honors College to complete Capstone Experience/Thesis Project. Bowling Green, KY.

Bolton, C. L., Brown, R. D. (2014) Adjusting Multiple Correlations for Regression Overfitting and Indirect Range Restriction. Paper to be presented at 2014 Annual Conference of the Society of Industrial/Organizational Psychologists. Honolulu.

FIELDS OF STUDY

Major Field: Psychology

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Every field of science depends on accurate measurement. Certain disciplines employ different methods of measurement, but the purpose of these tools is the same. Experiments with quantifiable results are necessary to determine whether a given hypothesis was supported. A given hypothesis is one possible explanation among many plausible explanations without a way to determine whether it conforms to reality. The importance of accurate measures is manifest for all fields of science, especially psychology, the study of human behavior.

One area of psychology that is highly dependent upon the accurate measurement of human behavior is industrial/organizational psychology. Industrial/organizational psychology, known as I/O psychology, is the study of human behavior in the work place. Measurement is the central component of many important functions in the field, including personnel selection, performance appraisal, and the assessment of training effectiveness.

Personnel Selection

For most hiring situations, there are more applicants than there are job openings. Thus, some method must be used to choose a subset of people to be hired from the applicant pool. In order to be fair for all the applicants and to maximize the value of the selection process to the organization, it is important that these hiring decisions are made based on accurate data.

Criterion-related validity studies are commonly employed to gather validity evidence in support of employment selection techniques. One way to conduct a criterion-related validity study is to utilize a predictive design. A criterion-related validity study with a predictive design proceeds as follows: (a) job-applicants complete a selection test, (b) some subset of these applicants are hired, (c) job performance (the criterion) is measured once the newly hired people have learned the job and developed a familiarity with the organization, and (d) scores on the selection test and the measure of job performance are correlated. A significant correlation is seen as evidence that the test is a good predictor of job performance.

One problem with a criterion-related validity study is range restriction. Range restriction is the truncation of one or both variables in a study and results in a sample correlation that underestimates the population validity (Guion, 1998). Range restriction can occur when prospective employees are chosen in a non-random manner from the applicant sample. There are two types of range restriction, direct and indirect. Direct range restriction occurs when applicants are selected for employment based on their scores on the test being validated (called the experimental predictor). Indirect range restriction occurs when selection decisions are made, not on the test being validated, but on a different test (called the operational predictor), a test which has a less than perfect correlation with the experimental predictor. The magnitude of the experimental-operational predictor intercorrelation determines the impact of the range restriction. The ideal scenario involves an operational predictor that is uncorrelated with the experimental predictor; no range restriction effects are experienced in such a situation. Range restriction, whether direct or indirect, can be avoided altogether by making selection decisions by randomly hiring applicants; however, this solution is often not practical for the organization.

According to the SIOP Principles (SIOP, 2003), sample correlations should be adjusted in order to "…obtain as unbiased an estimate as possible of the validity of the predictor in the population in which it is used" (p. 19). Thus, when the sample correlation is lowered by range restriction, it is desirable to apply certain correction equations to increase the correlation, offsetting the damage caused by range restriction. Thorndike (1949) provides the following formula to correct for the effects of indirect range restriction:

$$R_{xy} = \frac{r_{xy} + r_{xz}r_{yz}\left(\frac{S_Z{}^2}{s_z{}^2} - 1\right)}{\sqrt{\left[1 + r_{xz}{}^2\left(\frac{S_Z{}^2}{s_z{}^2} - 1\right)\right]\left[1 + r_{yz}{}^2\left(\frac{S_Z{}^2}{s_z{}^2} - 1\right)\right]}} \qquad (1)$$

Where: R_{xy} is the unrestricted correlation between the experimental predictor (i.e., the test in question) and the criterion (measure of job performance).

r_{xy} is the restricted correlation between the experimental predictor and the criterion.

r_{yz} is the restricted correlation between the criterion and the operational predictor.

r_{xz} is the restricted correlation between the experimental and the operational predictors.

$S_Z{}^2$ is the unrestricted variance of the operational predictor.

$s_z{}^2$ is the restricted variance of the operational predictor.

Several studies have been conducted testing the accuracy of this and other formulas. Lee, Miller, and Graham (1982) found that correlations were closer to the true values when corrected for range restriction effects. Other studies (e.g., Brown, Stout, Dalessio, & Crosby, 1988) have shown some overestimation in the correction. The accuracy of the correction likely depends on whether the regression assumptions of linearity and homoscedasticity are strictly supported in the population.

Criterion-related validity studies can accommodate more than one predictor to determine personnel decisions. Multiple regression analysis is frequently used to combine these multiple predictors to obtain the best possible prediction of performance. In a multiple regression analysis, it is important to decide how to weigh each score used so that there is the least possible error of prediction. Optimal weights, derived from a multiple regression analysis, are frequently seen as being the best; however, this leads to a new problem: the weights are optimized for the sample in which they are derived and will not predict as well when applied to future samples (Pedhazur, 1997). When applied to future samples, this sample-specific optimization (also called overfitting) leads to a reduction in validity, known as shrinkage. Thus, the multiple correlation coefficient obtained in the first sample is an upwardly biased estimate of the operational validity of this set of predictors.

The SIOP Principles (SIOP, 2003) also state that "…estimates of the validity of a composite battery developed on the basis of a regression equation should be adjusted using the appropriate shrinkage formula or be cross-validated on another sample" (p. 20). There are two types of methods proposed for estimating the most accurate multiple correlation: empirical and formula based. The empirical method involves applying regression weights from one sample to another sample drawn from the same population. The estimated cross-validated multiple correlation is the correlation between the predicted and actual criterion scores in the second sample (Mosier, 1951).

The empirical cross-validation method has a problem of its own: obtaining the second sample. This can be done by actually collecting a second sample of data, but such an approach is extremely labor intensive and can take an unreasonable amount of time. A more common method involves splitting a single sample into two subsamples. However, larger sample sizes lead to more stable results (Pedhazur, 1997). Thus, when splitting a sample there is greater sampling error, leading to higher overfitting in the first sample. Because of these reasons, researchers (e.g., Cascio, 1991) have suggested that empirical methods are less efficient than the formula methods.

The formula-based methods have been examined for accuracy (Raju, Bilgic, Edwards, & Fleer, 1999). It was found that using formulae estimators instead of empirical cross-validation can be done without significant reduction in the accuracy of the estimate of the shrunken correlation. Raju et al. (1999) looked at numerous available formulae and found that Burket's (1964) formula yielded the best results. Burket's equation is defined as follows:

$$P_{cv} = \frac{NR^2 - k}{R(N-k)} \tag{2}$$

Where:

P_{cv} is the estimated *unsquared* population cross-validated multiple correlation.

N is the sample size.

R^2 is the sample based squared multiple correlation coefficient.

k is the number of predictors.

Thus, there are two factors which affect the sample correlation computed in a predictive criterion-related validity study (with range restriction) for a battery of selection tests optimally weighted in a multiple regression analysis. One factor, range restriction, artificially lowers the sample correlation. The other factor, regression overfitting, artificially raises the sample correlation. It cannot be safely assumed that the inflated correlation from overfitting and the decreased correlation from range restriction will cancel each other out. Researchers should correct for both factors (SIOP, 2003). However, the effects of correcting for both factors using equations designed to correct for only one factor are unknown. Thus, it is worth investigating how to combine the two correction formulae. This investigation raises another question: Does the order of the correction equations matter? To date, there has been no empirical effort to investigate *indirect* range restriction and regression overfitting. The present study employs a Monte Carlo design to investigate the effects of both artifacts and the value of the two adjustment equations, as well as the order in which they are applied.

CHAPTER 2

METHOD

Overview

Data consisting of 1,000,000 cases, each with scores on one criterion variable and five predictor variables were generated with SAS version 9.2; all variables were normally distributed. These cases represent the population for the study. Bivariate correlations between the criterion and the experimental predictor variables were set as follows: r_{x1y} = .30, r_{x2y} = .30, r_{x3y} = .40, r_{x4y} =.40. The correlation between the operational predictor and the criterion variable was set as .30 for all conditions. The sample size of the selected group was 150 for all conditions.

Experimental Conditions

Three variables were manipulated. First, experimental predictor intercorrelations were set at either moderate (.40) or low (.20) levels. Second, in order to induce range restriction, people were hired at two distinct selection ratios of .10 and .30, meaning that either 10% or 30% of the applicants were selected. These values represent realistic conditions. Third, the bivariate correlation between the operational predictor and the individual experimental predictors was set at .30 or .50. For all three variables (experimental predictor intercorrelation, selection ratio, and experimental/operational predictor intercorrelation), these values were chosen to represent realistic conditions.

Because multiple predictors are being combined in a multiple regression analysis, the correlation between the experimental predictor composite and the operational predictor (e.g., R_{xz}) should also be considered. The various combinations described resulted in the following values for R_{xz}: .47 (predictor intercorrelation = .20, bivariate operational predictor correlations = .30), .79 (predictor intercorrelation = .20, bivariate operational predictor correlations = .50), .41 (predictor intercorrelation = .40, bivariate operational predictor correlations = .30), and .67 (predictor intercorrelation = .40, bivariate operational predictor correlations = .50).

For each of the eight conditions, samples were randomly drawn from the population 1000 times. Analyses were performed on each, and the results were averaged across these 1000 replications.

CHAPTER 3

PROCEDURE

Both Artifacts.

Analyses of the combined effects of regression overfitting and indirect range restriction were performed as follows. First, a sample of applicants was randomly selected from the population of one million cases. Because this study is designed to investigate indirect range restriction, hiring decisions were made top-down on the operational predictor. Experimental predictor scores for the hired cases were then optimally weighted in an OLS multiple regression analysis to yield the squared multiple correlation coefficient. This correlation served as the starting point for the various correlation adjustments as it is influenced by the combined effects of indirect range restriction and regression overfitting. This correlation was then adjusted using Burket's estimator of the population cross-validated correlation (Equation 2) and Thorndike's indirect range restriction adjustment (Equation 1) to yield a correlation free from the effects of these artifacts. This adjusted correlation, once squared, was then compared to the squared population cross-validated correlation (explained below) to determine the accuracy of these adjustment procedures. The effects of these dual adjustments were further investigated by varying the order of the adjustments to determine whether there is an order effect.

Single Artifact Baselines.

To further investigate the efficacy of these adjustment equations, each applicant sample was also selected or validated in a manner that generated only one artifact: indirect range restriction or regression overfitting. These single artifact conditions serve as baseline conditions for the dual artifact correction and act as a guide to the relative efficiency of the dual artifact corrections. Both baseline conditions were computed within the same applicant sample as the dual artifact condition of the experiment. For the first baseline condition, only regression overfitting was utilized to calculate the resultant correlation. This condition proceeded as follows. Within each applicant sample, applicants were selected via a random variable so that range restriction would not affect the results. Predictor scores for the selected cases were then optimally weighted via a multiple regression analysis. The resultant squared multiple correlation was then adjusted for regression overfitting only.

For the second baseline condition, only the effects of indirect range restriction were induced. Applicants were selected from within the applicant sample top-down on scores on the operational predictor. Experimental predictor scores were then unit weighted (using means and standard deviations from the entire applicant sample). The unit weighted composite score was then correlated with the criterion to yield the sample correlation. The sample correlation was then corrected with the indirect range restriction adjustment equation.

Population Cross-Validated Correlations.

To evaluate the effectiveness of these procedures (including baseline conditions), the estimated validities were compared to the population cross-validated squared correlations. Population cross-validated correlations were computed as follows. For the dual artifact condition as well as the first baseline condition (i.e., regression overfitting only), the sample regression equations were applied to all 1,000,000 cases in the population. The resultant composite scores were then correlated with the criterion variable, yielding the population cross-validated correlations. For the second baseline condition (i.e., range restriction only), the sample means and standard deviations were used to compute unit weighted composite scores for all 1,000,000 cases in the population. The correlation between this composite score and the criterion variable served as the population correlation.

The squared adjusted sample correlations were subtracted from the squared population cross validated correlations to yield the dependent variables for this study: bias (the signed difference) and squared bias. The mean of these values gives the average accuracy (bias) and an index of the variability of the bias (squared bias). As mentioned, results were averaged across 1000 replications.

CHAPTER 4

RESULTS

Mean bias and squared bias were computed across all results taken from the 1,000 samples. In some cases (about five percent of the time), correlations between the operational predictor and criterion variable or experimental predictor composite were negative. These cases were removed and were not considered for the remainder of the study. Simulations were run until there were 1,000 sets of results with positive correlations. The samples yielding negative correlations were not analyzed further because using the indirect range restriction correction equation with a negative correlation *lowers* the adjusted correlation. In addition, a negative sample-based correlation involving the operational predictor is a sign of a deviant sample (the researcher expects a positive correlation in population due to the probable hypotheses about the various predictor variables) and is not likely to be considered for a range restriction adjustment. It is recommended that indirect range restriction corrections should not be performed when the predictor intercorrelation is negative until this problem is better understood.

Table 1 lists the mean bias and mean squared bias for the baseline conditions as well each order of the dual artifact corrections for all eight experimental conditions. Table 2 lists Cohen’s *d* for bias and squared bias across all eight experimental conditions for the following comparisons: correcting both artifacts versus not correcting for either in samples affected by both, correcting for both artifacts versus regression overfitting baseline, correcting for both artifacts versus range restriction baseline, and correcting for both artifacts in varying order. Cohen's *d* statistics were computed to index effect size and make for easy comparison of the relative accuracy of the various correction procedures (i.e., comparing each condition to each other instead of to the cross-validated population correlations). In addition, inferential significance tests were not computed due to the nature of Monte Carlo experimentation (i.e., the abnormally large number of samples possible in Monte Carlo studies can result in significant results in which observed differences are trivial in magnitude).

Using Cohen's *d* for the data sets in Table 1 allows for an easy comparison of the experimental conditions to analyze whether these correlation adjustments are beneficial when the statistical artifacts are present. When both artifacts are present but the correlation is not adjusted the greatest level of bias is present (in seven of the eight conditions). However, not correcting for these artifacts also resulted in some of the lowest levels of squared bias in any of the comparisons. Thus, given these results, it seems that using the correction equations improves bias while worsening squared bias. This is most likely due to the indirect range restriction corrections. Correlations corrected for this artifact (whether the baseline or dual artifact conditions) displayed greater levels of squared bias than the baseline condition without range restriction. In summary, the increased accuracy (i.e., reduced bias) associated with correcting for indirect range restriction, no matter what the case, comes as a price: increased levels of squared bias.

A second question to consider is whether the correcting for both artifacts results in a less accurate estimate of the population cross-validated correlation than correcting for just one of these artifacts, when only one artifact has affected the results. Comparison of the dual artifact adjustment to the single artifact baseline conditions shows that adjustments for the effects of both of these artifacts are seldom less accurate than adjustments for just one of these artifacts (i.e., the baseline conditions). To this point, Cohen's *d* for bias never exceeds .2 for any of the eight conditions when a dual correction is compared to a single artifact baseline condition. Results are different for squared bias; squared bias for the comparison of adjustments for both artifacts to the regression overfitting baseline ranges from .19 to .67 (Cohen's *d*). As discussed above, this result is likely inherent to the indirect range restriction correction (note that when a dual correction is compared to the indirect range restriction baseline condition, Cohen's *d* for squared bias is less than .1 in every experimental condition). In summary, correcting for both artifacts, when present, is no worse in terms of bias and squared bias than a correction for just indirect range restriction when range restriction is the only artifact present.

The final issue investigated is whether the order of the corrections makes a difference in accuracy of the adjustment. Tables 1 and 2 show that the order of the corrections does not appear to make a difference in either bias or squared bias and Cohen's *d* values never exceed .1 for either statistic. Thus, researchers are free to choose the order to adjust for the effects of regression overfitting and indirect range restriction without consequence as the order of the adjustment to these does not affect the accuracy of the adjustment.

CHAPTER 5

DISCUSSION

Range restriction artificially lowers the sample correlation, while regression overfitting artificially raises it. Because it cannot be safely assumed these factors will cancel each other out, researchers must correct for both factors. This study was conducted in order to determine which method leads to the more accurate corrected sample correlation.

Using a Monte Carlo study, which allows the researcher to compare the sample correlation to the population cross-validated correlation, a population consisting of 1,000,000 cases was generated. Samples were drawn from the population and variables were manipulated to generate and correct for both statistical artifacts as well as single artifact situations. The single artifact baselines were used to further investigate the efficacy of these adjustment equations. Analyses were averaged across 1000 replications each. The difference between these averages and the population cross-validated correlations, known as bias and squared bias, were also computed. From studying the results, there are three main points: using the correction equations reduces bias while inflating squared bias as compared to uncorrected values affected by range restriction and regression overfitting; adjustments for the effects of both artifacts are about as accurate as an adjustment for just one of these artifacts when only indirect range restriction is present; and the order of corrections does not appear to make a difference.

There are four conclusions that can be drawn from the results of this study. First, when a correlation is affected by both indirect range restriction and regression overfitting, using both correction methods yields an adjusted sample correlation closer to the population cross-validated correlation than the unadjusted correlation. Second, the indirect range restriction correction improves the bias but adversely affects the squared bias of the correlation. This increase in squared bias is found whether the correction occurs alone or along with the regression overfitting correction. Third, correlations corrected for the combined effects of both artifacts exhibit bias that is close to that observed in correlations affected by, and corrected for, indirect range restriction alone. The fourth and final conclusion is that the order of correction has no effect on the accuracy of the corrections. Thus, researchers are free to correct for either artifact first without any negative effects.

Future research should investigate the accuracy of these multiple corrections in alternate conditions. The present study used varied selection ratios and predictor intercorrelations but held constant the validities of the operational and experimental predictors. Varying the values of these variables could lead to different conclusions. Also, the greater levels of squared bias observed with the indirect range restriction correction should be further investigated. Future researchers should explore the use of a different indirect range restriction correction equation (Thorndike, 1949, lists an alternate version).

The implications of this study to future researchers and practitioners are clear. Researchers and practitioners can and should use corrections for both regression overfitting and indirect range restriction where these artifacts are present to properly estimate of the validity of the test battery. Furthermore, these adjustments can be performed in either order without adversely affecting the accuracy of the resultant validity estimate.

REFERENCES

Brown, S. H., Stout, J. D., Dalessio, A. T., & Crosby M. M. (1988). Stability of validity indices through test score ranges. *Journal of Applied Psychology, 73,* 736-742.

Burket, G. R. (1964). A study of reduced rank models for multiple prediction. *Psychometrika Monograph Supplement*, No. 12.

Cascio, W.F. (1991). Validation and use of individual differences measures. *Applied Psychology in Personnel Management* (4th ed.), 149-170. Englewood Cliffs, NJ: Prentice-Hall.

Guion, R. M. (1998). *Assessment, Measurement, and Prediction for Personnel Decisions.* Mahwah, NJ: Erlbaum.

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.

Mosier, C. (1951). The need and means of cross validation. I. Problems and designs of cross-validation. *Educational and Psychological Measurement, 11,* 5-11.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace.

Raju, N., Bilgic, R., Edwards, J., & Fleer, P. (1999). Accuracy of population validity and cross-validity estimation: An empirical comparison of formula-based, traditional empirical, and equal weights procedures. *Applied Psychological Measurement, 23,* 99-125.

Society for Industrial Organizational Psychology, Inc. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Author.

Thorndike, R. L. (1949). *Personnel selection.* New York: Wiley.

APPENDIX

Table 1

Mean Bias and Mean Squared Bias for Given Validity Coefficient Corrections

	Bias	Squared Bias
Condition 1		
Regression Overfitting; Regression Overfitting (Baseline)	.0042	.0042
Range Restriction; Range Restriction (Baseline)	-.0020	.0053
Both Artifacts; No Adjustments	.0109	.0036
Both Artifacts; Range Restriction then Regression Overfitting	.0033	.0055
Both Artifacts; Regression Overfitting then Range Restriction	.0035	.0054
Condition 2		
Regression Overfitting; Regression Overfitting (Baseline)	.0057	.0044
Range Restriction; Range Restriction (Baseline)	-.0165	.0093
Both Artifacts; No Adjustments	-.0079	.0038
Both Artifacts; Range Restriction then Regression Overfitting	.0008	.0101
Both Artifacts; Regression Overfitting then Range Restriction	-.0067	.0094
Condition 3		
Regression Overfitting; Regression Overfitting (Baseline)	.0032	.0038
Range Restriction; Range Restriction (Baseline)	-.0020	.0046
Both Artifacts; No Adjustments	.0045	.0033

Both Artifacts; Range Restriction then Regression Overfitting	.0026	.0051
Both Artifacts; Regression Overfitting then Range Restriction	.0041	.0051
Condition 4		
Regression Overfitting; Regression Overfitting (Baseline)	.0079	.0038
Range Restriction; Range Restriction (Baseline)	-.0059	.0060
Both Artifacts; No Adjustments	.0172	.0034
Both Artifacts; Range Restriction then Regression Overfitting	.0024	.0064
Both Artifacts; Regression Overfitting then Range Restriction	.0015	.0061
Condition 5		
Regression Overfitting; Regression Overfitting (Baseline)	.0052	.0044
Range Restriction; Range Restriction (Baseline)	-.0093	.0063
Both Artifacts; No Adjustments	.0165	.0038
Both Artifacts; Range Restriction then Regression Overfitting	-.0031	.0064
Both Artifacts; Regression Overfitting then Range Restriction	-.0030	.0063
Condition 6		
Regression Overfitting; Regression Overfitting (Baseline)	.0072	.0045
Range Restriction; Range Restriction (Baseline)	-.0298	.0127
Both Artifacts; No Adjustments	-.0084	.0035
Both Artifacts; Range Restriction then Regression Overfitting	-.0113	.0130
Both Artifacts; Regression Overfitting then Range Restriction	-.0200	.0123
Condition 7		
Regression Overfitting; Regression Overfitting (Baseline)	.0066	.0041

Range Restriction; Range Restriction (Baseline)	-.0061	.0061
Both Artifacts; No Adjustments	.0093	.0037
Both Artifacts; Range Restriction then Regression Overfitting	-.0009	.0065
Both Artifacts; Regression Overfitting then Range Restriction	.0005	.0064
Condition 8		
Regression Overfitting; Regression Overfitting (Baseline)	.0038	.0039
Range Restriction; Range Restriction (Baseline)	-.0114	.0082
Both Artifacts; No Adjustments	.0260	.0037
Both Artifacts; Range Restriction then Regression Overfitting	-.0037	.0085
Both Artifacts; Regression Overfitting then Range Restriction	-.0047	.0080

Note. The descriptions in each line are formatted as follows: “Artifact(s) affecting the correlation; adjustments performed on the correlation.” Sample size in each condition was 150 (i.e., 150 people were hired). Selection ratio was .33 for Conditions 1-4 and was .10 for Conditions 5-8. Experimental predictor intercorrelations were .20 for Conditions 1, 2, 5, and 6 and were .40 for Conditions 3, 4, 7, and 8. Correlation between operational predictor and optimally weighted experimental predictor was .30 for Conditions 1, 3, 5, and 7 and was .50 for Conditions 2, 4, 6, and 8.

Table 2

Cohen's d for Mean Bias and Mean Squared Bias for Comparisons of Given Validity Coefficient Corrections

	Bias	Squared
Bias		
Condition 1		
Both Adjustments vs. No Adjustments	.111	.290
Both Adjustments vs. Regression Overfitting Baseline	.013	.189
Both Adjustments vs. Range Restriction Baseline	.072	.033
Both Adjustments vs. Opposite Order	.003	.010
Condition 2		
Both Adjustments vs. No Adjustments	.015	.578
Both Adjustments vs. Regression Overfitting Baseline	.058	.561
Both Adjustments vs. Range Restriction Baseline	.178	.059
Both Adjustments vs. Opposite Order	.076	.048
Condition 3		
Both Adjustments vs. No Adjustments	.006	.288
Both Adjustments vs. Regression Overfitting Baseline	.009	.205
Both Adjustments vs. Range Restriction Baseline	.066	.062
Both Adjustments vs. Opposite Order	.021	.003
Condition 4		
Both Adjustments vs. No Adjustments	.232	.431
Both Adjustments vs. Regression Overfitting Baseline	.077	.368
Both Adjustments vs. Range Restriction Baseline	.106	.049
Both Adjustments vs. Opposite Order	.012	.033
Condition 5		
Both Adjustments vs. No Adjustments	.278	.331
Both Adjustments vs. Regression Overfitting Baseline	.114	.268
Both Adjustments vs. Range Restriction Baseline	.078	.013
Both Adjustments vs. Opposite Order	.002	.018
Condition 6		
Both Adjustments vs. No Adjustments	.133	.727
Both Adjustments vs. Regression Overfitting Baseline	.199	.672
Both Adjustments vs. Range Restriction Baseline	.167	.021
Both Adjustments vs. Opposite Order	.078	.044
Condition 7		
Both Adjustments vs. No Adjustments	.125	.363
Both Adjustments vs. Regression Overfitting Baseline	.105	.310
Both Adjustments vs. Range Restriction Baseline	.065	.045
Both Adjustments vs. Opposite Order	.019	.013

Condition 8		
Both Adjustments vs. No Adjustments	.414	.503
Both Adjustments vs. Regression Overfitting Baseline	.095	.506
Both Adjustments vs. Range Restriction Baseline	.085	.029
Both Adjustments vs. Opposite Order	.011	.046

Note. The numbers in each column are unsigned Cohen's *d*. Sample size in each condition was 150 (i.e., 150 people were hired). Selection ratio was .33 for Conditions 1-4 and was .10 for Conditions 5-8. Experimental predictor intercorrelations were .20 for Conditions 1, 2, 5, and 6 and were .40 for Conditions 3, 4, 7, and 8. Correlation between operational predictor and optimally weighted experimental predictor was .30 for Conditions 1, 3, 5, and 7 and was .50 for Conditions 2, 4, 6, and 8.