
Correcting Correlations for Measurement Error and Range Restriction

Correlations are mathematically affected by two common problems, measurement error and range restriction. Correcting for these effects will increase the correlation.

1. Correlations are indicators of relations among pairs of variables and the magnitude of a relation is reduced to the extent that those variables contain random responses. This is caused by measurement error – that is, each variable is measured with some amount of measurement error or random variation. The more measurement error present in a set of scores, the more the correlation is reduced (not because of a lack of relationship but because more of the variance is random and randomness does not correlate with anything).
2. Correlations are based on covariance, which assumes variation in scores. If measures do not vary, there is no way to assess the strength of a relationship. To the extent that two or more groups result in scores with different levels of variance, the correlation will subsequently be affected such that the group with lower variance will result in lower correlations, all else equal. Another way to think about this is range restriction – lower variance is range restriction, which reduces correlations.

Correction for Measurement Error

$$r_{XY}^* = \frac{r_{XY}}{\sqrt{r_{XX}} \sqrt{r_{YY}}} \quad \text{where} \quad r_{XY}^* = \text{the corrected correlation between } X \text{ and } Y;$$

r_{XY} = the original correlation;

r_{xx} = the reliability of scores X ;

r_{yy} = the reliability of scores Y .

Correction for Range Restriction

$$r_{XY}^* = \frac{S_X r_{XY}}{\sqrt{S_x^2 (1 - r_{XY}^2) + S_X^2 r_{XY}^2}} \quad \text{where} \quad r_{XY}^* = \text{the corrected correlation};$$

r_{XY} = the original correlation;

S_X = the restricted group standard deviation;

S_x = the unrestricted group standard deviation.

Note. When correcting for both measurement error and range restriction, the correction for measurement error should be done first, followed by the correction for range restriction. This order is only appropriate when the range restriction (when based on purposive selection of participants) is not based on one of the variables composing the correlation.