Chapter 20

Accommodating Measurement Errors

Joop J. Hox
Department of Methodology & Statistics, Utrecht University

Most if not all measures employed in surveys contain some amount of measurement error. One reason is that surveys contain many measures that consist of a single survey question, and if the questionnaire contains multi-item scales they typically use a small number of questions. As a result, the reliability of these instruments is not high. This chapter discusses the consequences of low reliability on the estimates produced with such instruments. Typically, low variability increases the variance of the measurements, and bivariate relations are underestimated. When multi-item scales are used, adjustments for measurement error can be made by using structural equation modeling and including the separate items in an explicit measurement model. When single items are used, adjustments can be made only if information is available on their reliability, such as test-retest correlations from another study. A very different approach to adjusting for measurement errors is to impute for each respondent a set of plausible values, analogous to procedures used in multiple imputation of missing values.

GLOSSARY OF KEY POINTS

Attenuation. The reduction of the estimated bivariate relationship between variables when one or both of these is measured with low reliability.
Classical true score theory. A measurement model that decomposes each measurement into a true score and an error component. The main object is to estimate the proportion of true score variance in a measure, which is defined as its reliability. The limitations of true score theory have led to the development of *item response theory*.
Item response theory (IRT). Statistical measurement models that assume a mathematical model for the probability that a given subject will respond correctly (positively, agree to) to a given question.
Latent variable. A characteristic that can not be observed or measured directly. It is hypothesized to exist in order to explain observed variables. Also called factor.

Measurement error. Lack of measurement precision due to flaws in the measurement instrument. If measurement errors are random they decrease the *reliability*, if they are systematic they decrease the *validity*.
Model fit. How closely the model-implied data match the observed data. In SEM model fit can be tested using a formal chi-square test (*p*-values >.5 indicate good fit), or it can be evaluated using model fit indices such as CFI (>0.9 indicates good fit) or RMSEA (<.05 indicates good fit).
Plausible values. Estimated latent score values drawn at random from a conditional distribution, given the responses to the items and a set of background variables (conditioning variables).
Reliability. Absence of random measurement errors.
Structural equation model. A multivariate model describing the relationships between multiple observed and/or latent variables. Often referred to as SEM.
Validity. Absence of systematic measurement errors.