The versatility of G-theory for exploring and controlling measurement error

Sandra Johnson
Assessment Europe, Scotland

Since publication of *The Dependability of Behavioral Measurements*, generalizability theory (G-theory) has moved on. In particular, the versatility of the methodology has become increasingly appreciated by growing numbers of practitioners in diverse domains of application. Jean Cardinet and his colleagues were instrumental in this process, illustrating how symmetry in the underlying ANOVA (analysis of variance) model can fruitfully be exploited in practice.

With ANOVA , we partition the total variance in a set of data into component parts, each attributable to a particular variance source (test items, psychological attributes, markers, schools, interaction between markers and items, etc), and quantify their relative contributions. With G-theory, we use the resulting information to estimate measurement error and calculate generalizability coefficients; "What if?" analyses allow us to *predict* reliability should the measurement procedure be adjusted for a future application (eg. increasing test length, number of markers, or whatever).

G-theory enables us to address more complex and more authentic measurement applications than classical test theory was developed to do. Indeed, classical test theory is in effect a special case of G-theory, while G-theory is itself a special case of multi-level modelling. Links with item response modelling are now also being explored and clarified. The presentation overviews the essentials of G-theory, using examples of application to illustrate points, and briefly considers the links between it and the other dominant modelling approaches in educational measurement.