Structural Analysis of Subjective Categorical Data
Karl Christoph Klauer, William H. Batchelder

A general approach to analysis of subjective categorical data is considered, in which agreement matrices of two or more raters are directly expressed in terms of error and agreement parameters. The method provides focused analyses of ratings from several raters for whom ratings have measurement error distributions that may induce bias in the evaluation of substantive questions of interest. Each rater's judgment process is modeled as a mixture of two components: an error variable that is unique for the rater in question as well as an agreement variable that operationalizes the "true" values of the units of observation. The statistical problems of identification, estimation, and testing of such measurement models are discussed. The general model is applied in several special cases. The most simple situation is that underlying Cohen's Kappa, where two raters place units into unordered categories. The model provides a generalization and systematization of the Kappa idea to correct for agreement by chance. In applications with typical research designs, including a between-subjects design and a mixed within-subjects, between subjects design, the model is shown to disentangle structural and measurement components of the observations, thereby controlling for possible confounding effects of systematic rater bias. Situations considered include the case of more than two raters as well as the case of ordered categories. The different analyses are illustrated by means of real data sets.