Stella Method

*[general introduction]*

*Cluster Analysis.*

First, we used the K-means clustering with MDS. Therefore, we…

*[hierarchical clustering?]*

For determining the number of clusters …

*Factor Analysis.*

For determining the number of factors we used …

*[maximum likelihood, pearson correlations, oblique promax falls nötig]*

The calculations were programmed in the open source software R 0.94.110 using the package *[mit Zitat einfügen].*

*Traditional Simulation study.*

We specified the model from the NEO-PI-R with five factors and eight variables each using the … command in R. To make the simulation comparable to the real-world model, estimates of main factor loading, factor intercorrelations and uniquenesses were obtained from the factor analysis of the norm data set and subsequently used for the simulations. *[simulation conditions]*

Additionally, we also applied different sample sizes (100, 150, 250, 300, 500 and 1000).

Samples were produced from a multivariate normal distribution with mean=0 and the given population covariance matrix (calculated from loadings, factor intercorrelations and uniqenesses).

*[specific simulation conditons for each method, methods used, number of conditons, way of determining the number of clusters]*

*Real World simulation.*

*The data set.* The NEO-PI-R is a personality inventory measuring personality in five major domains: Neuroticism, extraversion, openness to experience, agreeableness and conscientiousness. Each domain scale is divided into six facets and eight items operationalize each facet. Thus, the questionnaire consists of 240 items.

We chose the following 5 facets from each of the 5 factors: N1 (Anxietey), E2 (Gregariousness), O3 (Feelings), A4 (Compliance) and C5 (Self-discipline). As indicated in the manual of the NEO-PI-R the correlation matrix of items shows intermediate intercorrelations of items within facets (from .18 to 0.36) and low intercorrelations between factors (for more details see Ostendorf & Angleitner, 2004).

Factor loadings are between .23 and .79. Although these values may seem particularly low, they are rather typical of personality questionnaires. Peterson (2000) showed in a meta-analysis on factor loadings in exploratory factor analyses of questionnaire data, that average factor loading is .32 with 25% of the factor loadings being less than .23, and 25% greater than .37. For this study, the self-report form was used in which the participant has to provide self-report about typical behaviours or reactions on a five point Likert scale, ranging from 0=‘strongly disagree’ to 4=‘strongly agree’. Validity and reliability for all questionnaires was shown by Ostendorf and Angleitner (2004). Means, standard deviations and skewnesses of the items are shown in table 1.

The norm data set consists of 12003 participants of which 279 are clinical patients, which we did not include. Therefore, the final sample, we used, consisted of 11724 participants. The mean age of the final group was 29.92 ranging from 16 to 91 with 36% males and 64% females. Sum scores on facets were calculated for each subject, while subjects with at least one missing value per facet got an NA instead of a sum score and therefore were not include in further analyses for this facet.

*Samples.* Smaller samples of sizes … were drawn, with 1000 replications each, randomly with replacement out of the whole data set. The sampling with replacement was chosen because otherwise a bigger sample would have meant automatically a closer similarity to the population data set and therefore better results. By choosing sampling with replacement we intended to create fairer conditions between different sample sizes. For assessment of dimensionality in CA for each sample … . In EFA for each condition … . Then, for each condition we determined how often the same number of factors/clusters as in the population data set was suggested and calculated the mean of the number of factors/clusters to determine the biasedness of the method. For assignment of facets in EFA a PAF analysis with pearson correlations, maximum-likelihood estimation and oblique Promax rotation was conducted.

*[alles was noch rein muss]*

Thus, for CA we had ?x? different conditions for determining the number of factors as we compared … and ? for assignment of facets. For EFA we had a fully crossed ?x?x?-design, since we compared … .

*Cross validation with CFA*.

After specifying factor models with each of these methods we subsequently tested the specified model with a CFA. We then reported Qui-square values and AIC and BIC of each model. …