Dynamic Item- and Teststatistics
A shiny GUI for test development
Erik Sengewald,
Department ‘Psychological Research and Development’, German Federal Employment Agency, Nürnberg

Requirements of the department
Over 70 different tests for psychological diagnostic (e.g., cognitive ability, language, occupational interests and more) are managed by the department and new tests are developed continually. Standardized procedures may facilitate test development, thus a scientist can focus on project specific tasks and challenges. Reproducibility and ease of use have to be core features of a statistic tool to meet different levels of experience with R and allow others to collaborate in a concrete test development procedure.

Overview & Input
Introduction
The problem
This app was developed because I needed a dynamic tool for item selection. I was facing the following problem:

- a bunch of test items have to be combined to a test
- item characteristics vary with the set of items i choose for the final test
- test characteristics change dynamically with the selected items
- all those problems appear for different subgroups

App features
The app is able to do a lot of stuff:

- filter your data to make a clean dataset before analyzing
- set group variables to view results for every group
- item characteristics (difficulty, discrimination) as summary table, graphs and interactive graphs
- test-score distribution for the total group and every group level
- difference testing of the test-score (sum) between different group levels (including standardized differences)
- some analysis of reliability
- test of unidimensionality (1 PL, 2 PL)
- test of any other measurement model
- test for measurement invariance
- test for a specified structure equation model with the testscore (sum)

Output

Updating the analysis by:
- choosing or reordering items
- changing data input or filter data
- selecting subgroups or reorder them
- loading an existing report (loads the state of the former report, too)
- or changing a relevant option of a current analysis

Note: all examples are based on simulated data