# Final project assignment

Final project comprises of validation and analysis of measurement data. Provided dataset is inspired by paper analyzing rating of applicants to teacher positions, see https://doi.org/10.1371/journal.pone.0203002. In the data, the columns represent applicant ID, rater ID, internal status, and rating on the 9 subcomponents closer described in the paper.

You are also encouraged to use your own dataset – in such case, first submit proposal describing your data and methods. You will get confirmation or further suggestions within 5 working days.

Your final project is to be submitted electronically in PDF form. Commented R code for your analysis should be attached as well. Long tables or high number of figures (e.g. such as for item distractor plots) may be moved into supplemental material, or may be provided as PDF file generated from ShinyItemAnalysis R package. Your project should be written in Czech or in English and should contain following sections:

## Introduction (2 pts)

Provide 1–4 paragraphs of an introduction to the study. Introduce the topic providing at least 3 references. Why was the study undertaken? What was the research question or the purpose of the research?

## Methods (2 pts)

Describe the study design and the data, describe any manipulation with the data such as variable transformations or recoding, treatment of potential missing values, omitted observations (1 pt). Describe the analytical methods and models (1 pt).

## Results (12 pts)

Provide description, tables and figures with results. Each referenced Table or Figure should be numbered and should appear after the paragraph where it is cited. Suggested analysis should comprise following paragraphs:

- A consise descriptive analysis (2 pts)
  Provide descriptive table(s) and/or figures.
- Proofs of test reliability (2 pts)
  Provide an estimate of internal consistency. Assess test-retest reliability, if retest is available. Provide an estimate of inter-rater reliability if more raters are present. Additionally, you may also assess IRR by group, or you may use Spearman-Brown formula to discuss how many items/raters would be needed to reach certain level of reliability (e.g. 0.9, 0.8 or 0.7) for averaged rating.
- Proofs of test validity (2 pts)
  Analyze correlation structure. Provide proofs of criterion validity if criterion is available.
- Item analysis using traditional methods or regression models (2 pts)
  Include table of traditional item indices (difficulty, RIR, RIT, ULI, alphaDrop,..). Include distractor plots for all items. Comment on items with highest and lowest difficulty, items of low discrimination or items with inappropriate distractor plot. Provide explanation taking into account wording of those items.
- Item response theory model (2 pts)
  Provide reasoning for selected IRT model (based on data type or using comparison of more IRT models). Plot and discuss a Wright map. Plot item characteristic curves, item information curves, test information curve. Provide table of model parameter estimates and their standard errors, ability estimates and their relationship to traditional ability estimates based on total scores.
- Differential item functioning (2 pts)
  Provide reasoning for selected DIF detection method, or try more methods and compare results. Discuss any items detected as DIF, provide wording or explanation for DIF.

If any of the above parts is not appropriate for your data, you need to explain why. In such case, points will be split in remaining parts of the analysis. You may also include other methods covered in class if appropriate.

## Discussion (2 pts)

Provide discussion to your results in 1 – 4 paragraphs. What are the results and are they expectable? What recommendations do you have for the analyzed assessment instrument (in terms of removing/rewording items, increasing reliability or validity, etc.). For what other situations/data is your analysis relevant? What are the limitations to your study/analysis? What is your conclusion?

## References (2 pts)

Provide full references to all cited papers.

## Supplemental Materials (4 pts)

Attach other supplemental material, such as automatically generated report or other file with distractor plots. Attach commented R code for (parts of) your analysis.

*Bonus A (5 points):* Use your own dataset.

*Bonus B (5 points):* If your test comprises of 20 or more items, prepare an adaptive test using your item wording and parameter estimates of selected IRT model.

*Bonus C (5 points):* Simulate data best reflecting study design and results from article on teacher applicant ratings, or reflecting your study design and using parameter estimates of selected IRT model. Provide R code for your simulation and generated data. Such data may in future be used for simulation studies. Data structure should allow running all previous analyses in your supplemental R code.

*Due date:* Final project (in PDF format, together with any other supplementary material) is due at least 3 working days before the oral exam.