University of Zurich[UZH]

Differential Item Functioning
DIF-Analysis

Master Rasch Seminar 11 – 25.11.2020

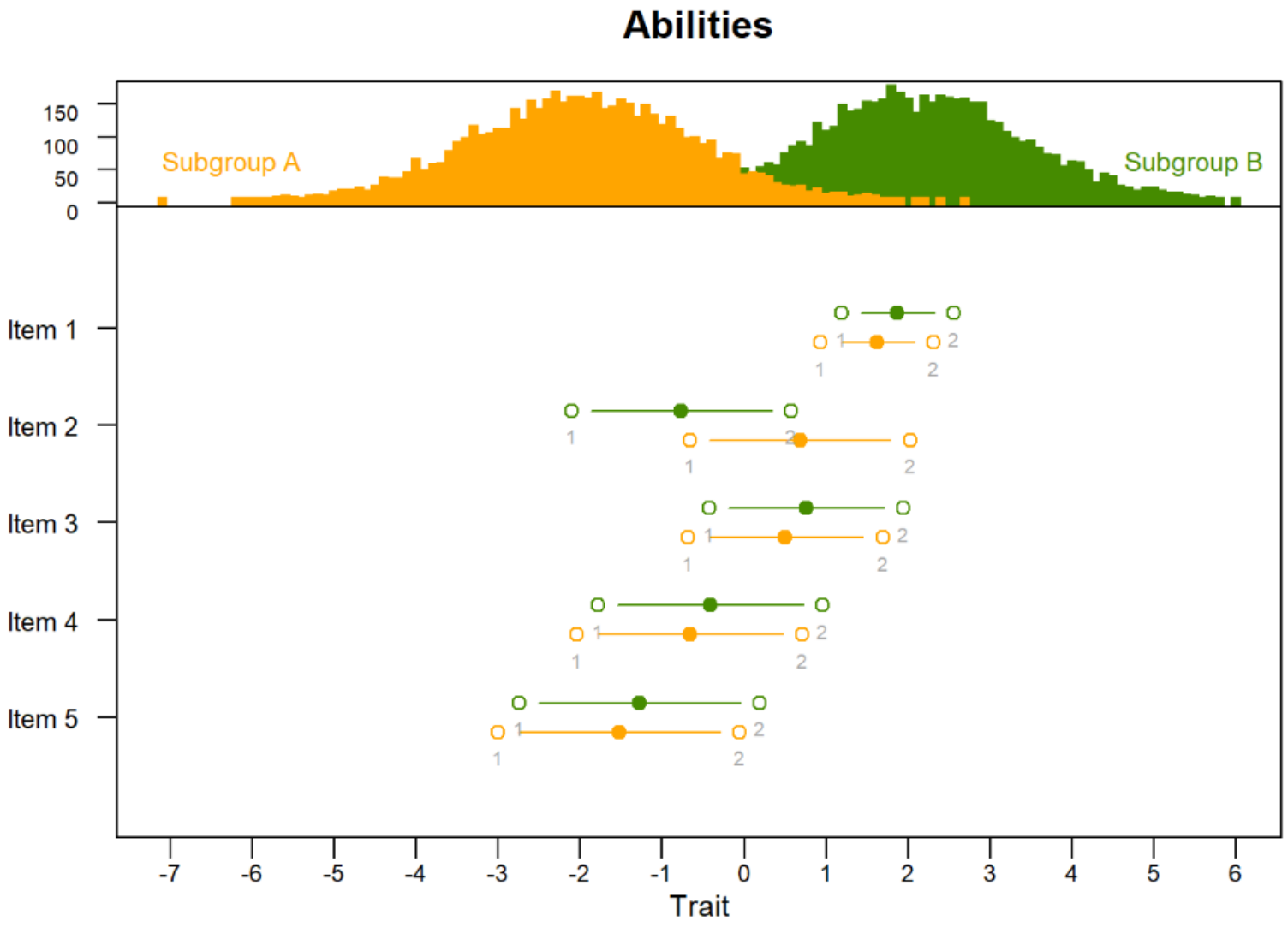
Carolina Fellinghauer : c.fellinghauer@psychologie.uzh.ch

Differential Item Functioning

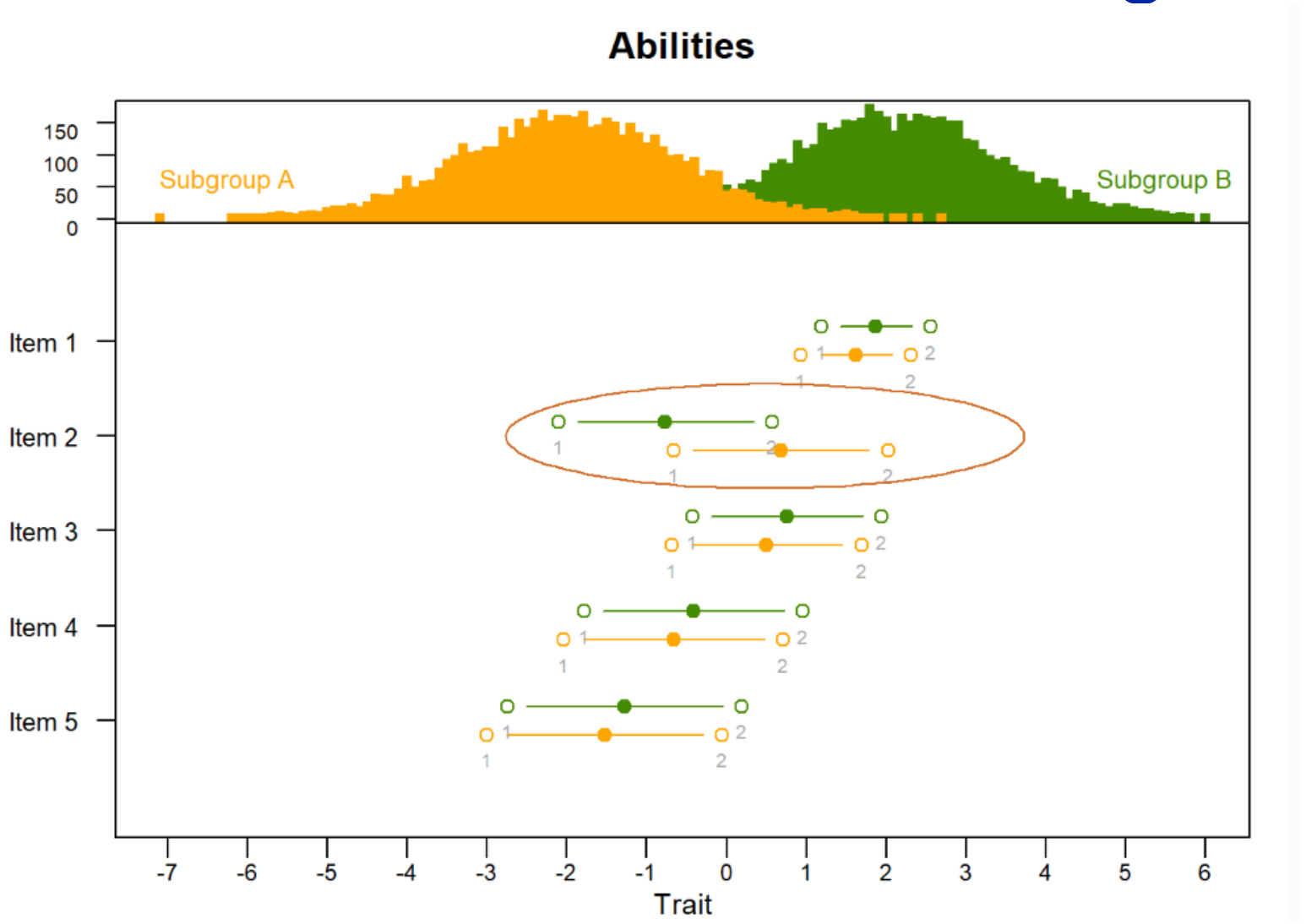
The Rasch model assumes the construct measured is valid across subgroups.

Differential item functioning tests if items are invariant across sample subgroups.

Differential Item Functioning

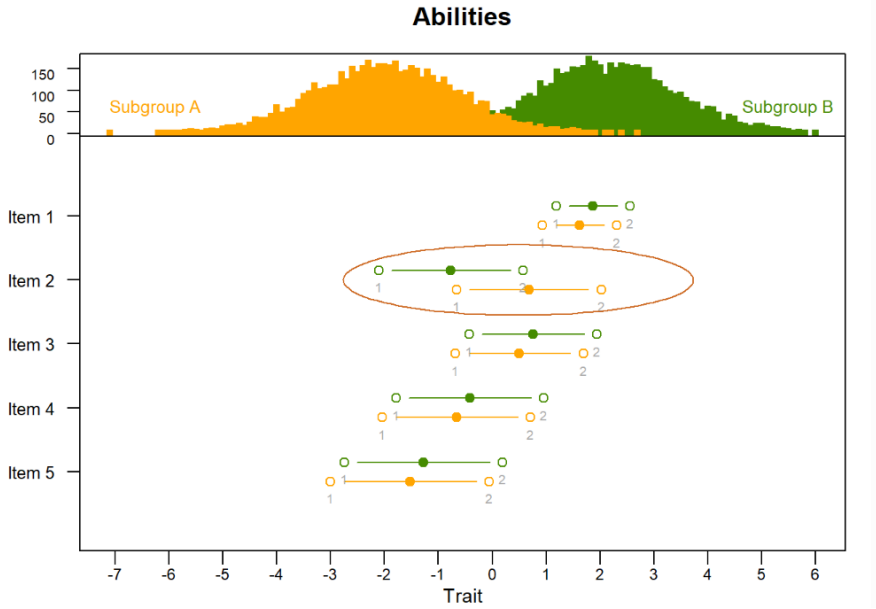


Differential Item Functioning



2) one item, the Item 2, is in different locations relative to the other items as a function of the subgroup.

Differential Item Functioning?

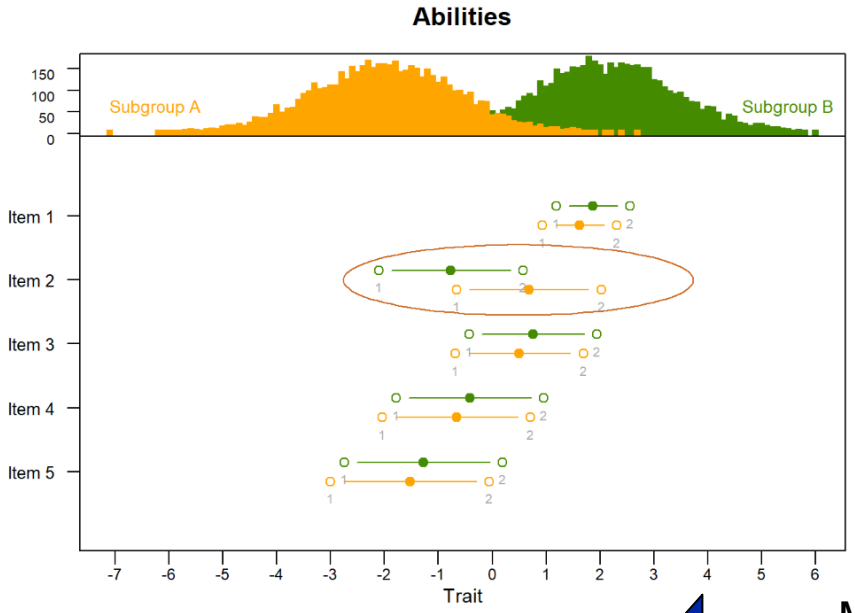


1) The ability of Subgroup A is lower than the ability of Subgroup B.

2) The difficulty of the item is similar for almost all items.

3) For a same level of ability, the difficulty of Item 2 differs across the Subgroup A and Subgroup B.

Differential Item Functioning?



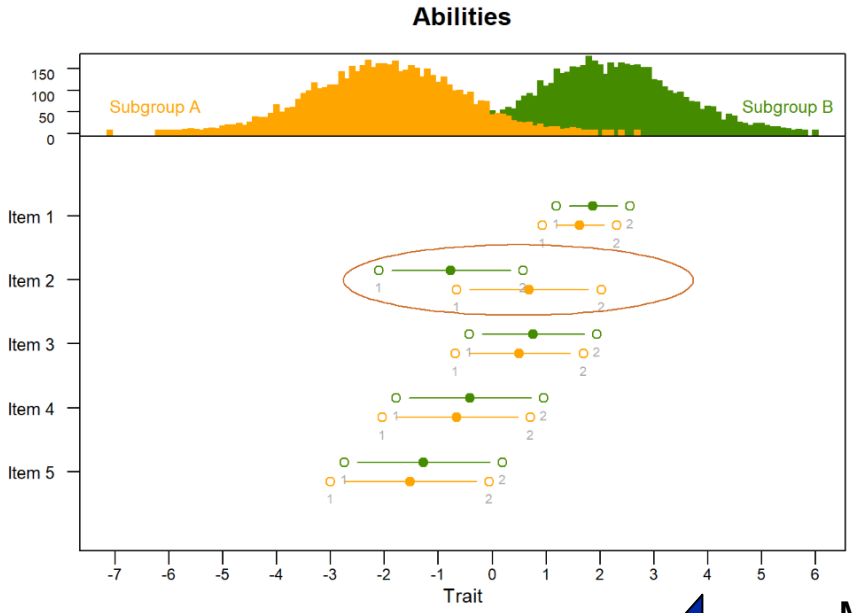
1) The ability of Subgroup A is lower than the ability of Subgroup B.

2) The difficulty of the item is similar for almost all items.

3) For a same level of ability, the difficulty of Item 2 differs across the Subgroup A and Subgroup B.

May not be a measurement bias or problem with the construct validity.

Differential Item Functioning?



1) The ability of Subgroup A is lower than the ability of Subgroup B.

2) The difficulty of the item is similar for almost all items.

3) For a same level of ability, the difficulty of Item 2 differs across the Subgroup A and Subgroup B.

May not be a measurement bias or problem with the construct validity.

DIF: Item 2 performs very differently in the two constructs.

DIF in Rasch Analysis

In Rasch analysis the residual matrix is tested for patterns that indicate different response types across subgroups.

One approach is a two way analysis of variance (ANOVA) of the residuals.

Two way because of: (1) a DIF variable (age, gender, language, survey year…) and (2) score groups and their interaction.

The score groups represent a division of the total scores into equal-sized score groups.

ID	Total Score	Score Group
1	1	
2	1	1
3	2	
4	3	
5	4	2
6	4	
7	5	
8	5	3
9	5	
10	6	
11	7	4
12	7	

Ideally the subgroup size should be between 30-50 persons.

A certain total score (example score = 2), is found only in one group.

Typically, the total score continuum would not be divided in to much more than 10 groups.

Differential Item Functioning
Uniform vs Non-Uniform DIF

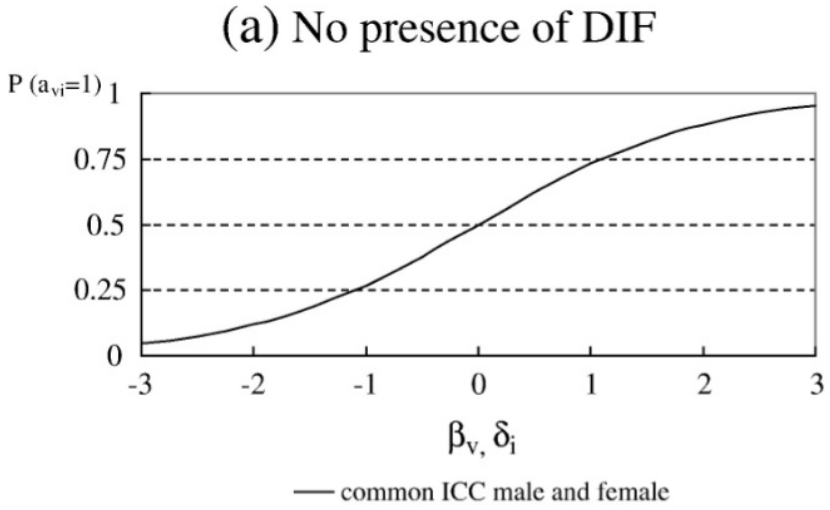
Uniform DIF:

Item difficulty estimates differ significantly across sample subgroups (age, gender, language, survey year, etc..)

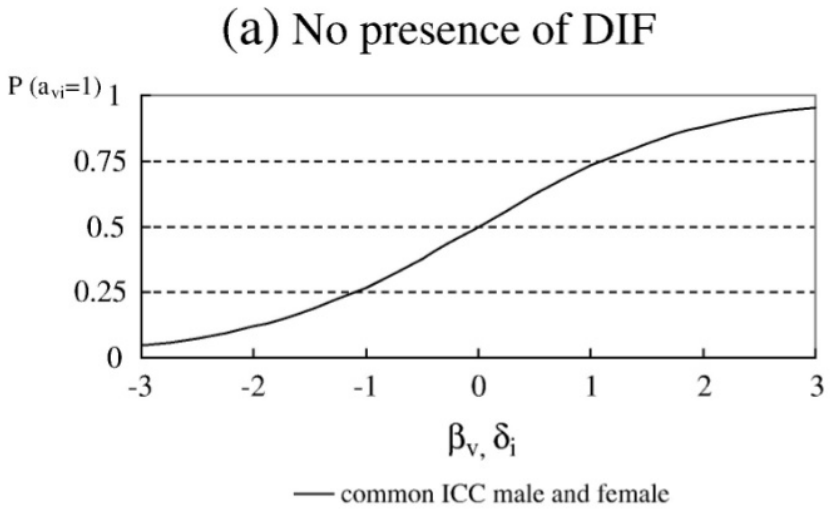
Non-Uniform DIF:

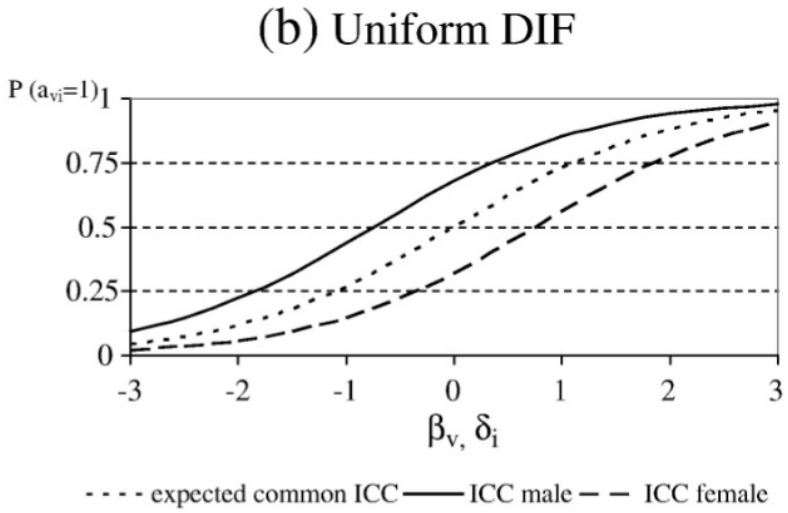
Item difficulty estimates differ significantly across sample subgroups and score level groups.

Differential Item Functioning

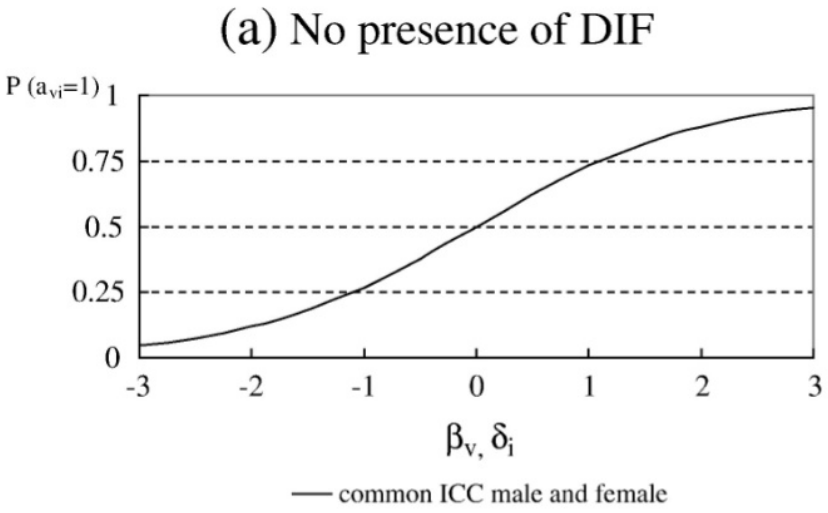


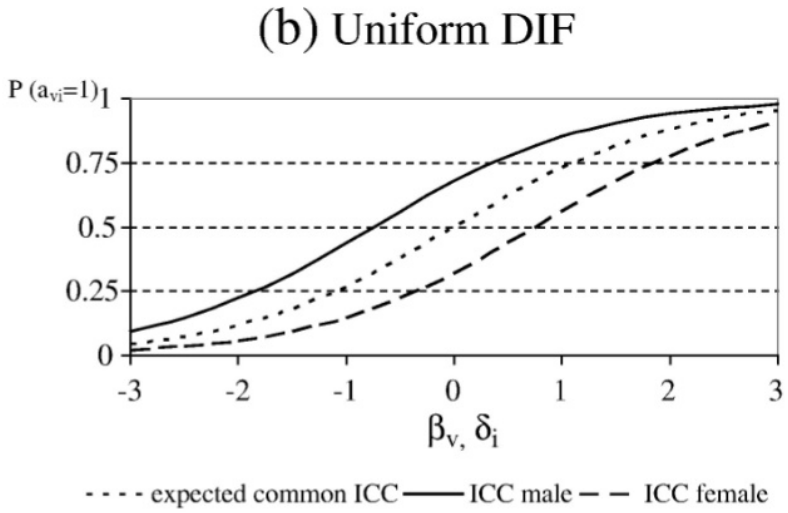
Differential Item Functioning

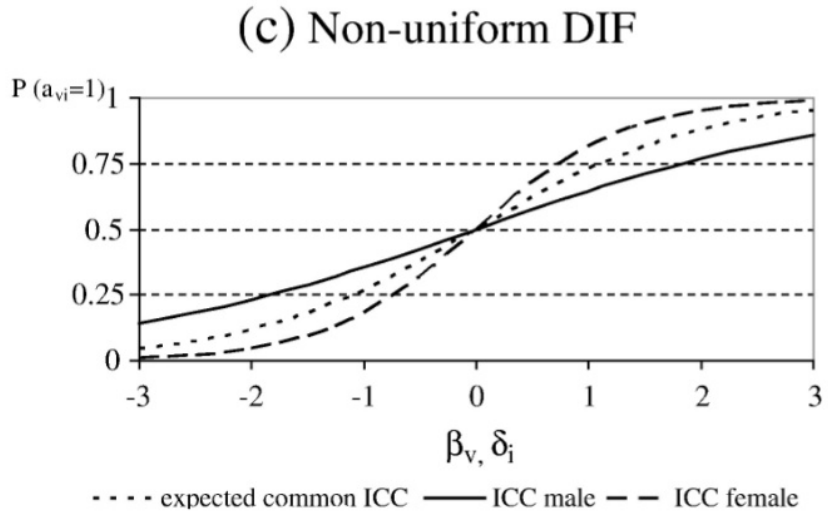




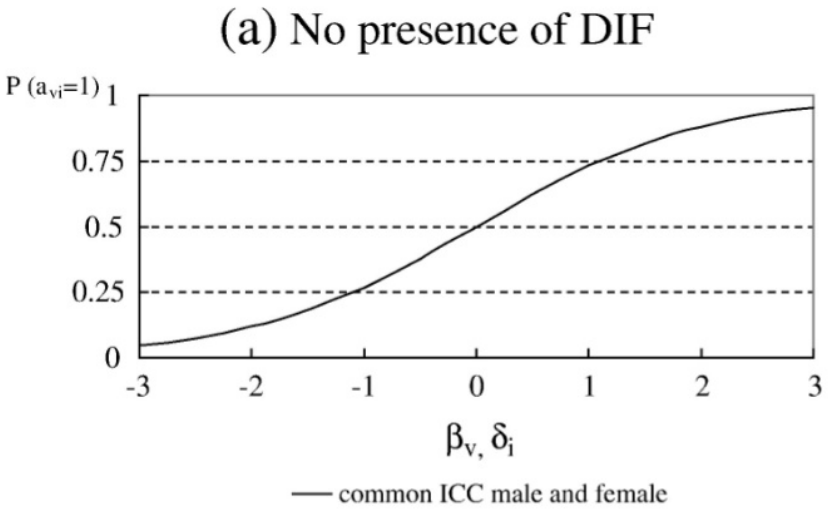
Differential Item Functioning

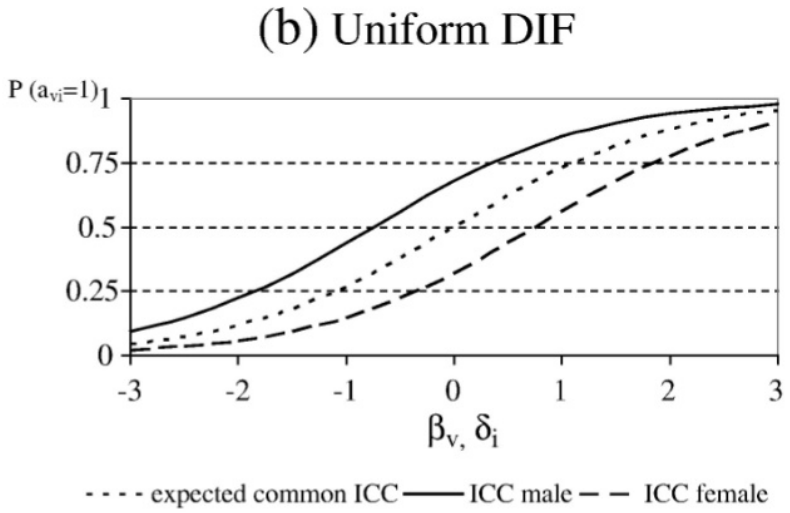


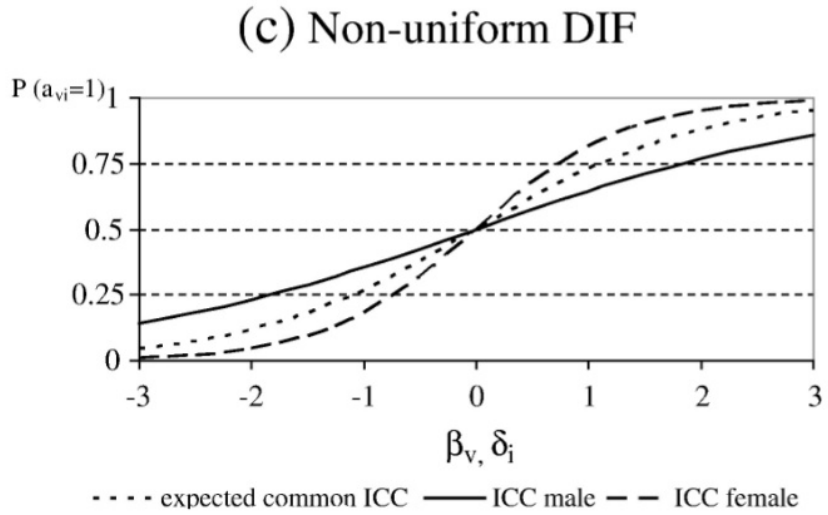


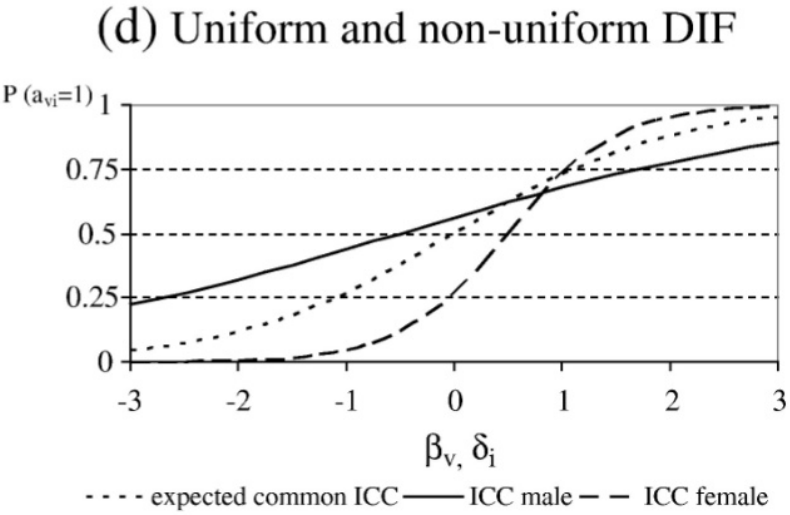


Differential Item Functioning

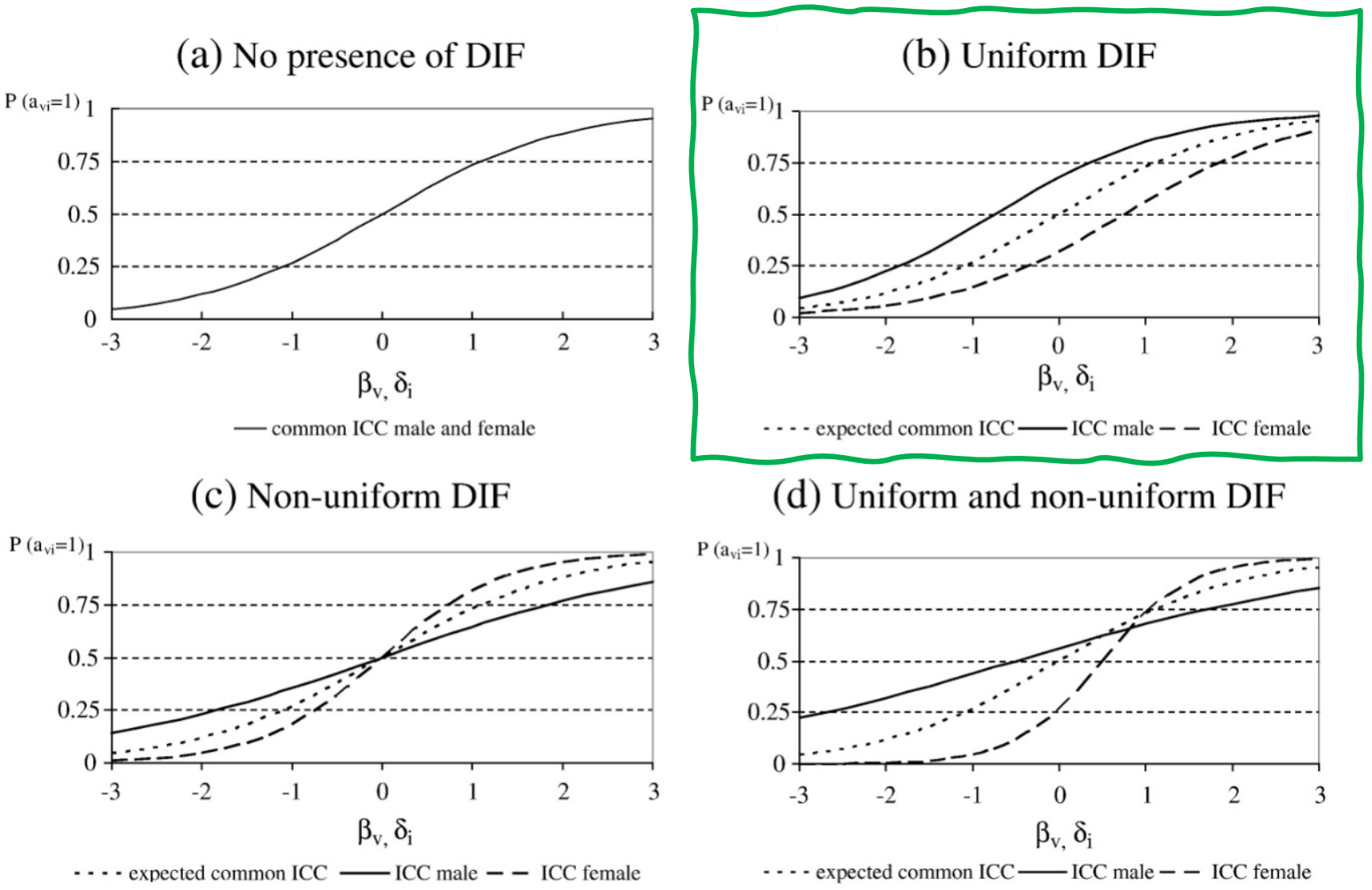






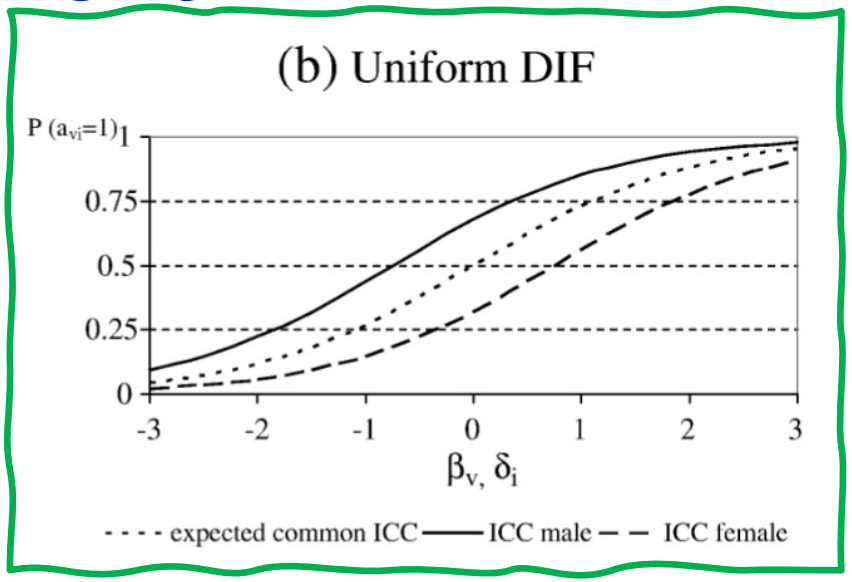


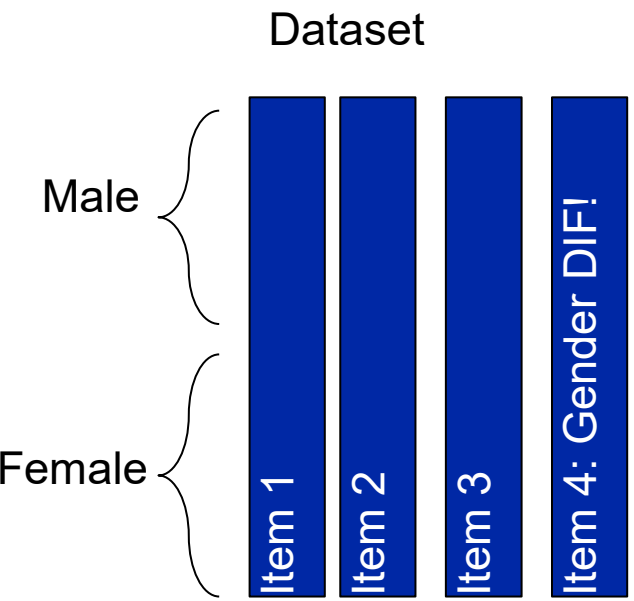
Differential Item Functioning



Differential Item Functioning Adjustment

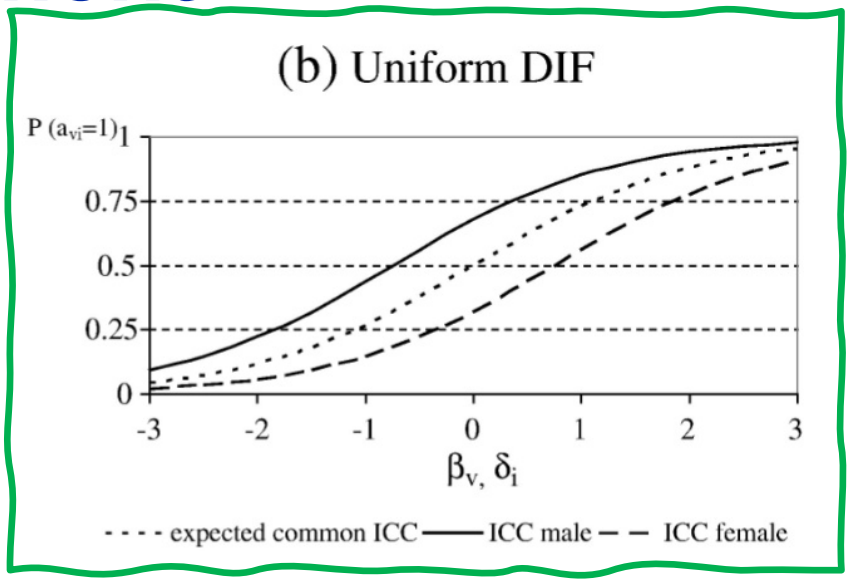
In presence of Uniform DIF, an approach to solve the DIF is item split.

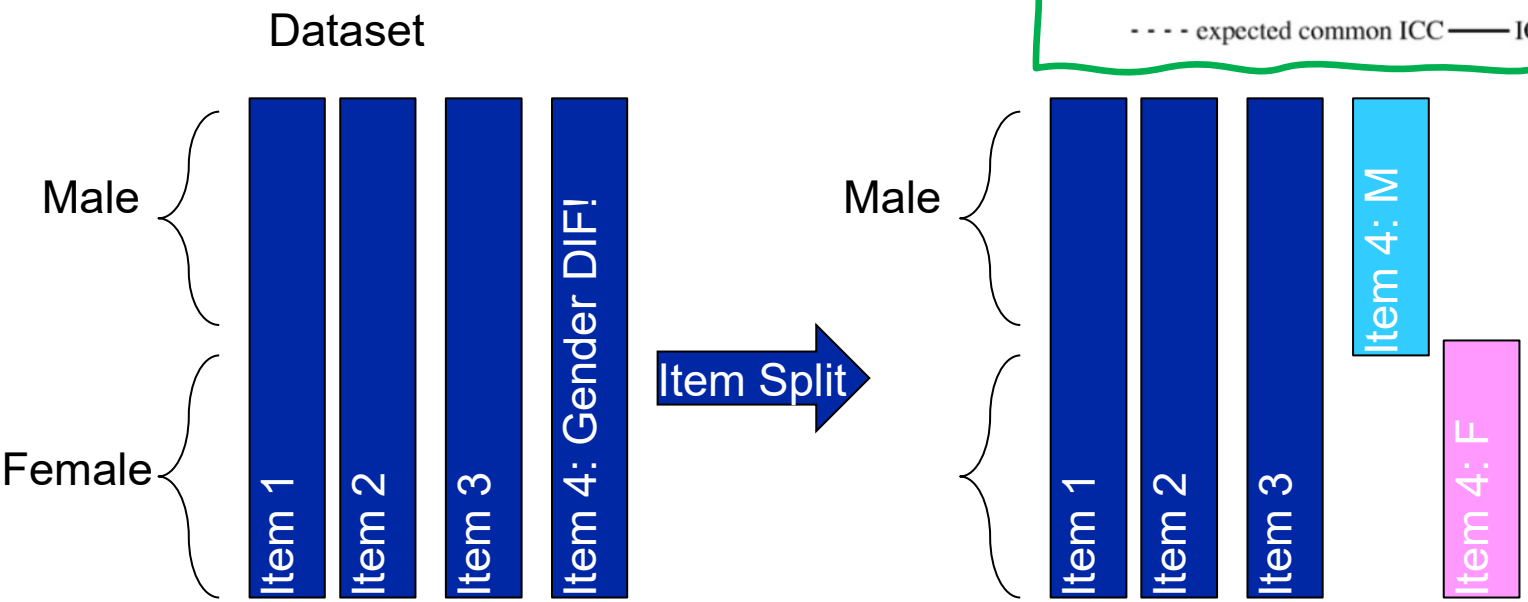




Differential Item Functioning Adjustment

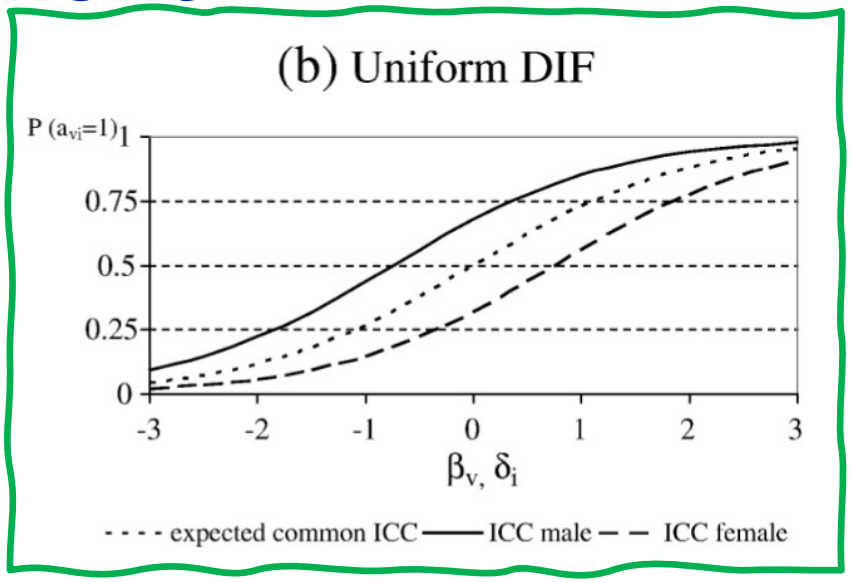
In presence of Uniform DIF, an approach to solve the DIF is item split.
Instead of one item, as many items as DIF subgroup level will enter the Rasch analysis.
This creates than adjusted subgroup specific metrics.

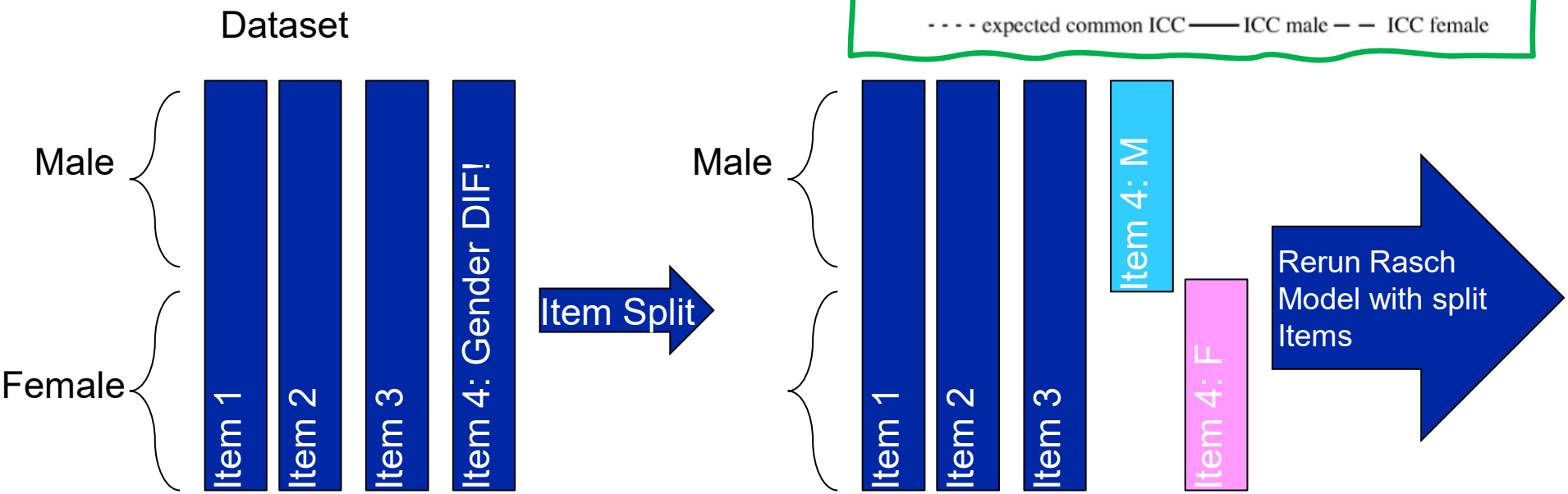




Differential Item Functioning Adjustment

In presence of Uniform DIF, an approach to solve the DIF is item split.
Instead of one item, as many items as DIF subgroup level will enter the Rasch analysis.
This creates than adjusted subgroup specific metrics.



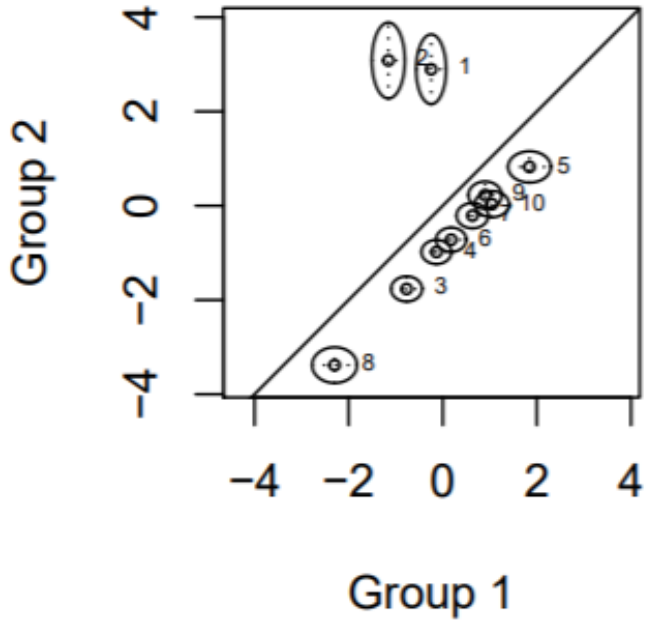


Artificial DIF

Item **splitting is done stepwise**, starting with the item with the highest DIF.

Given the PCM estimation approach which centers the item difficulties to zero. It is not unusal, that the residuals may show artificial DIF in some principally DIF-free items as a reaction to items with very high DIF.

Example: items free of DIF should be on the diagonal. In the plot below no item is on the diagonal. Stepwise solving of DIF starting with the two upper items, is likely to recenter the items and show absence of DIF in a second run.



Group 1 = ex. male
Group 2 = ex. female

Let’s go to R-Studio

Open the R-Script MS11_Rscript.r from the OLAT or the MS-Teams Course Materials.

Exercise

Run the same DIF analysis for Age group (< 30 < 45 < 60) and injury completeness. Detect uniform DIF and decide if you would need to split items or not.