World Health Organization

Introduction to Rasch Measurement

Carolina Fellinghauer, PhD, Dpl.Stat. External Consultant WHO
carolina.fellinghauer@who.int

16 4 2021

Psychometrics

Psychometrics is understood as the field of study concerned with the theory and technique of psychological measurement including the measurement of knowledge, abilities, attitudes, and personality traits. The field is primarily concerned with the study of differences between individuals.

It involves two major research tasks:

(i) the **construction** of instruments and procedures for measurement; and

(ii) the **development and refinement of theoretical approaches** to measurement

What is measurement?

Main ideas of measurement

What are the main ideas of measurement?

What is measurement?

Main ideas of measurement

What are the main ideas of measurement?

– **Objects** have **properties** that can be thought in terms of more or less, larger or smaller, stronger or weaker – example?

What is measurement?

Main ideas of measurement

What are the main ideas of measurement?

- **Objects** have **properties** that can be thought in terms of more or less, larger or smaller, stronger or weaker – example?
- This **property** can be measured through its **manifestation** – example?

What is measurement?

Main ideas of measurement

What are the main ideas of measurement?

- **Objects** have **properties** that can be thought in terms of more or less, larger or smaller, stronger or weaker – example?
- This **property** can be measured through its **manifestation** – example?
- This **manifestation** can be mapped onto a **scale** – example?

What is measurement?

Main ideas of measurement

What are the main ideas of measurement?

- **Objects** have **properties** that can be thought in terms of more or less, larger or smaller, stronger or weaker – example?
- This **property** can be measured through its **manifestation** – example?
- This **manifestation** can be mapped onto a **scale** – example?
- **Measurement** can have some **error** involved, and may not be perfectly precise – example?

Main ideas of measurement

Scale types with their properties according to Stanley Smith Stevens				
	Nominal scale	Ordinal scale	Interval scale	Ratio scale
Logical/ math operations × ÷	✗	✗	✗	✓
Logical/ math operations + −	✗	✗	✓	✓
Logical/ math operations > <	✗	✓	✓	✓
Logical/ math operations = ≠	✓	✓	✓	✓

Main ideas of measurement

Scale examples

What are some examples of each type of scale?

Main ideas of measurement

Scale examples

What are some examples of each type of scale?

– Nominal scale: 0=Apples, 1=Pears, 2=Bananas

Main ideas of measurement

Scale examples

What are some examples of each type of scale?

- Nominal scale: 0=Apples, 1=Pears, 2=Bananas
- Ordinal scale: 0=Never, 1=Sometimes, 2=Often

Main ideas of measurement

Scale examples

What are some examples of each type of scale?

- Nominal scale: 0=Apples, 1=Pears, 2=Bananas
- Ordinal scale: 0=Never, 1=Sometimes, 2=Often
- Interval scale: 0=0°C, 1=1°C, 2=2°C

Main ideas of measurement

Scale examples

What are some examples of each type of scale?

- Nominal scale: 0=Apples, 1=Pears, 2=Bananas
- Ordinal scale: 0=Never, 1=Sometimes, 2=Often
- Interval scale: 0=0°C, 1=1°C, 2=2°C
- Ratio scale: 0=0cm, 1=1cm, 2=2cm

Why is ordinal not good enough?

Pretend you are the CEO of an orange juice company 5000 oranges vs. 750 kg of oranges

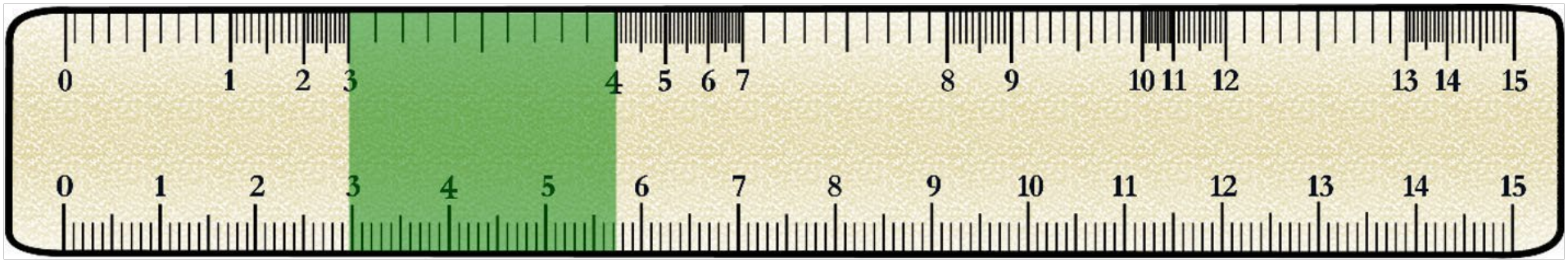
Measuring versus counting

1. Do you have difficulty engaging in vigorous activities, such as playing football or doing sport?
2. Do you have difficulty going outside of your home?
3. Do you have difficulty moving around inside your home?

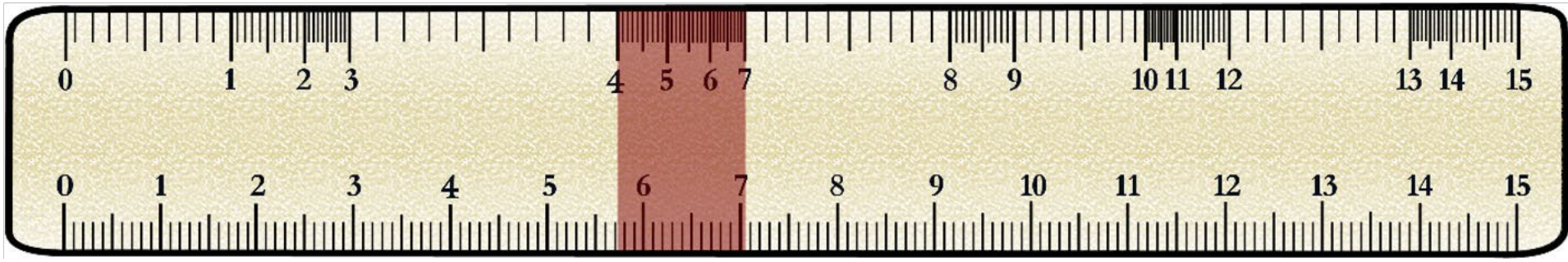
Not all questions are equally difficult!

Measurement invariance

Patient One



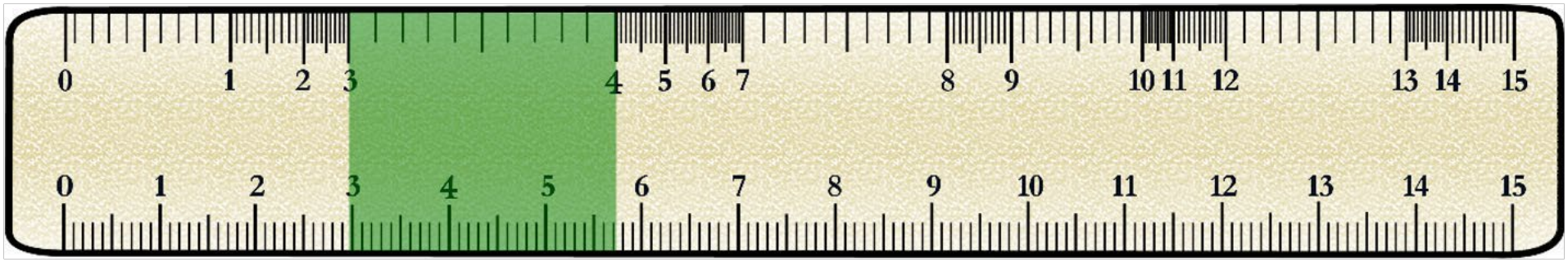
Patient Two



Measurement invariance

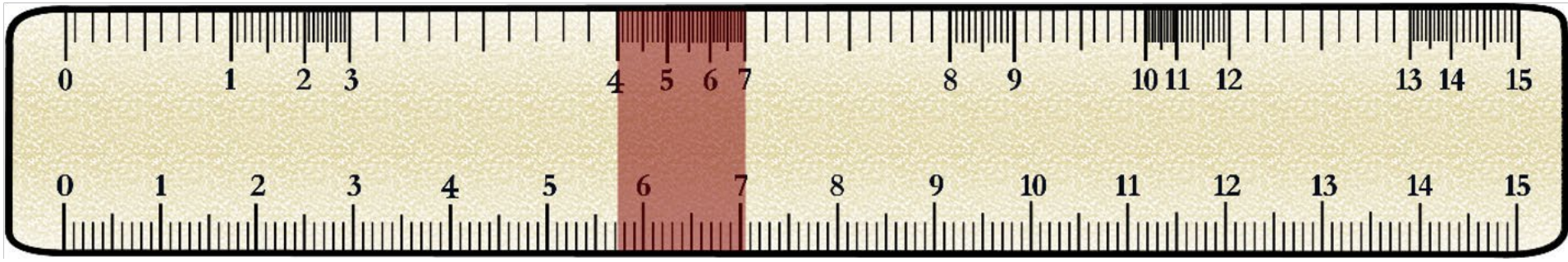
Patient One

+ 1



Patient Two

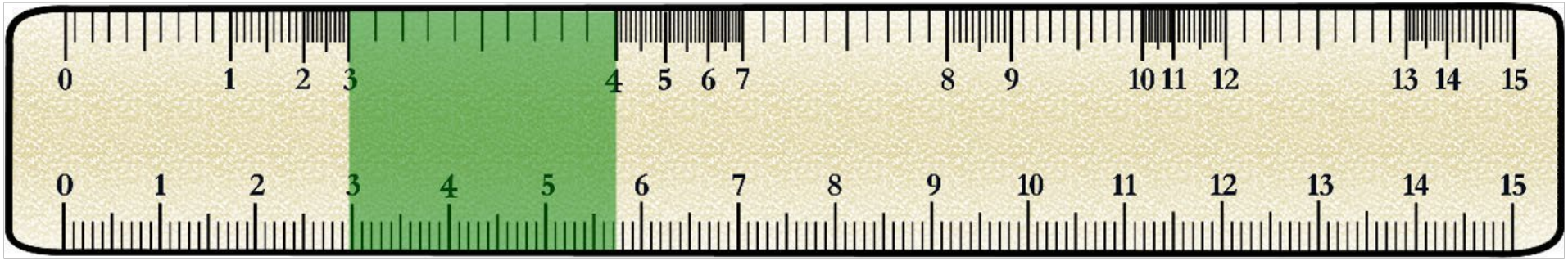
+3



Measurement invariance

Patient One

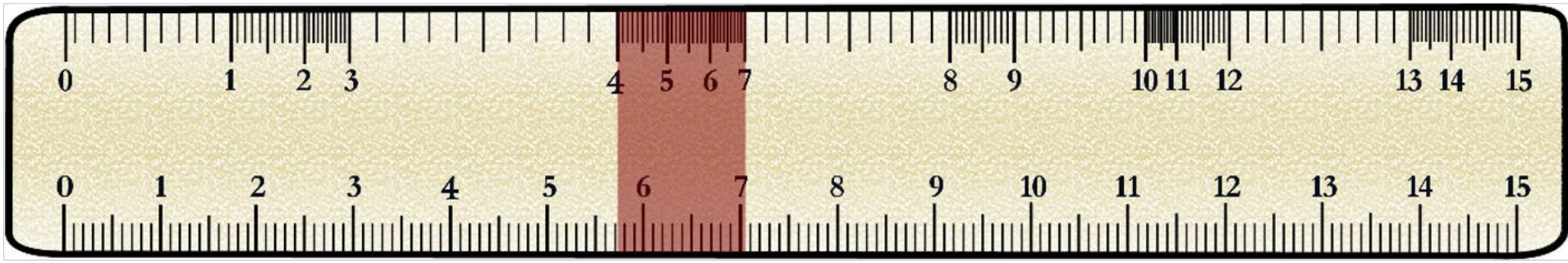
+ 1



+ 2.7

Patient Two

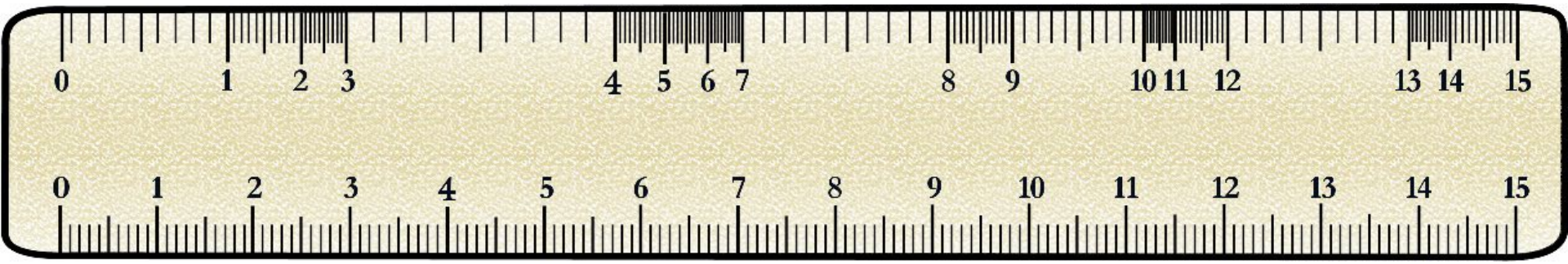
+3



+1.2

Our goal

Going from ordinal…

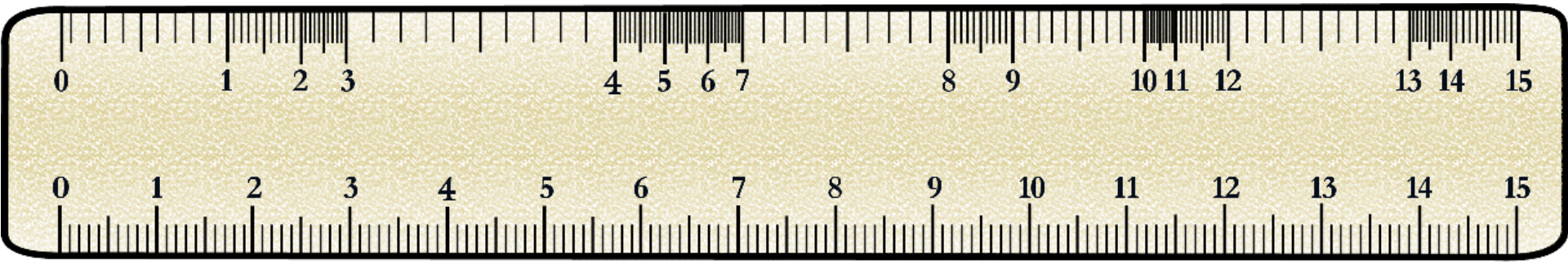


…to interval

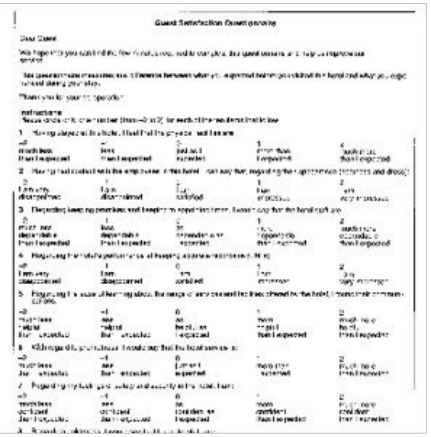
Misinference from Ordinal Scales[4]

"Ordinal scales of measurement
do not support the mathematical operations needed to
calculate means and standard deviations.

How to Make the Transition From the Top to the Bottom of the Ruler?



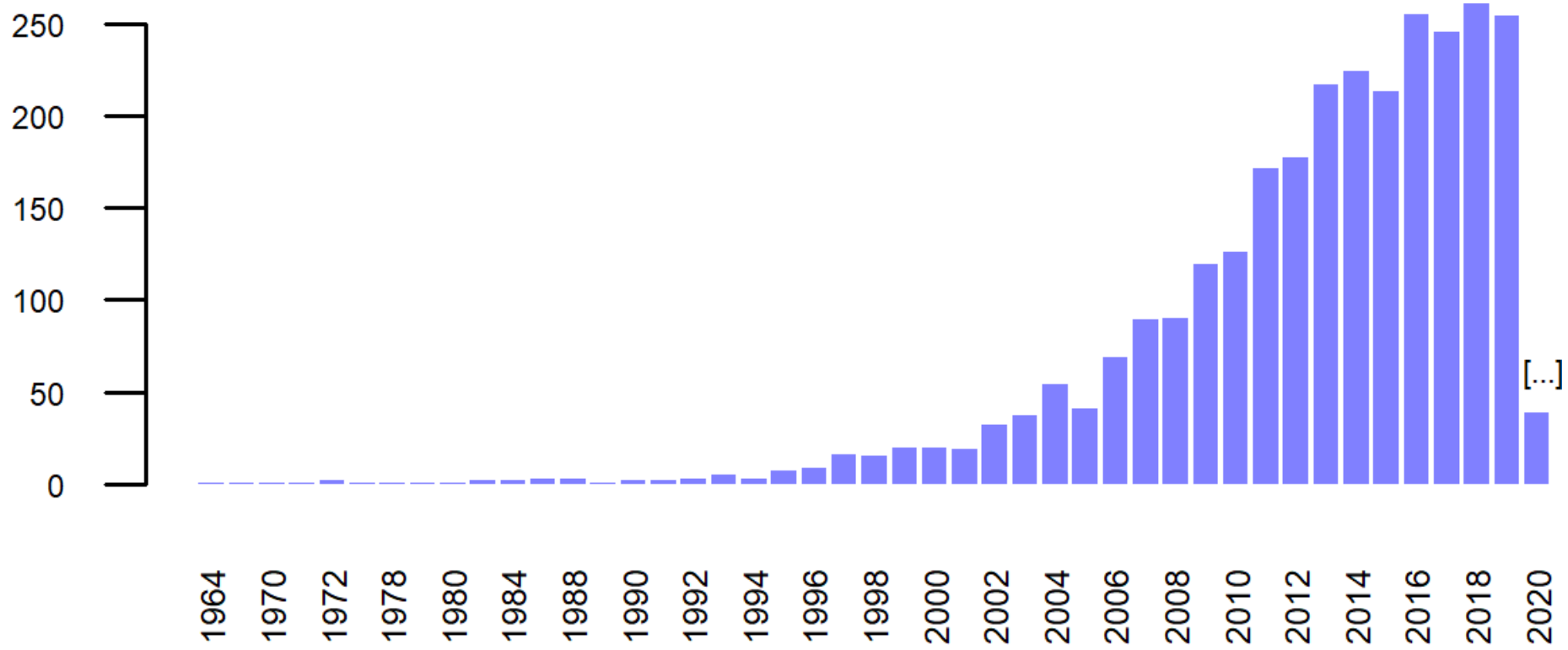
Rasch Analysis



$$p_i(\theta) = \frac{e^{(\theta-\delta_i)}}{1+e^{(\theta-\delta_i)}}$$

Georg Rasch (1901-1980)

Rasch-Publications per Year (N = 2855)



Pubmed search on Feb 2020:
https://pubmed.ncbi.nlm.nih.gov/
(Rasch Analysis) AND (Psychometry OR Test Theory OR Methods)

Rasch Analysis: What is that?

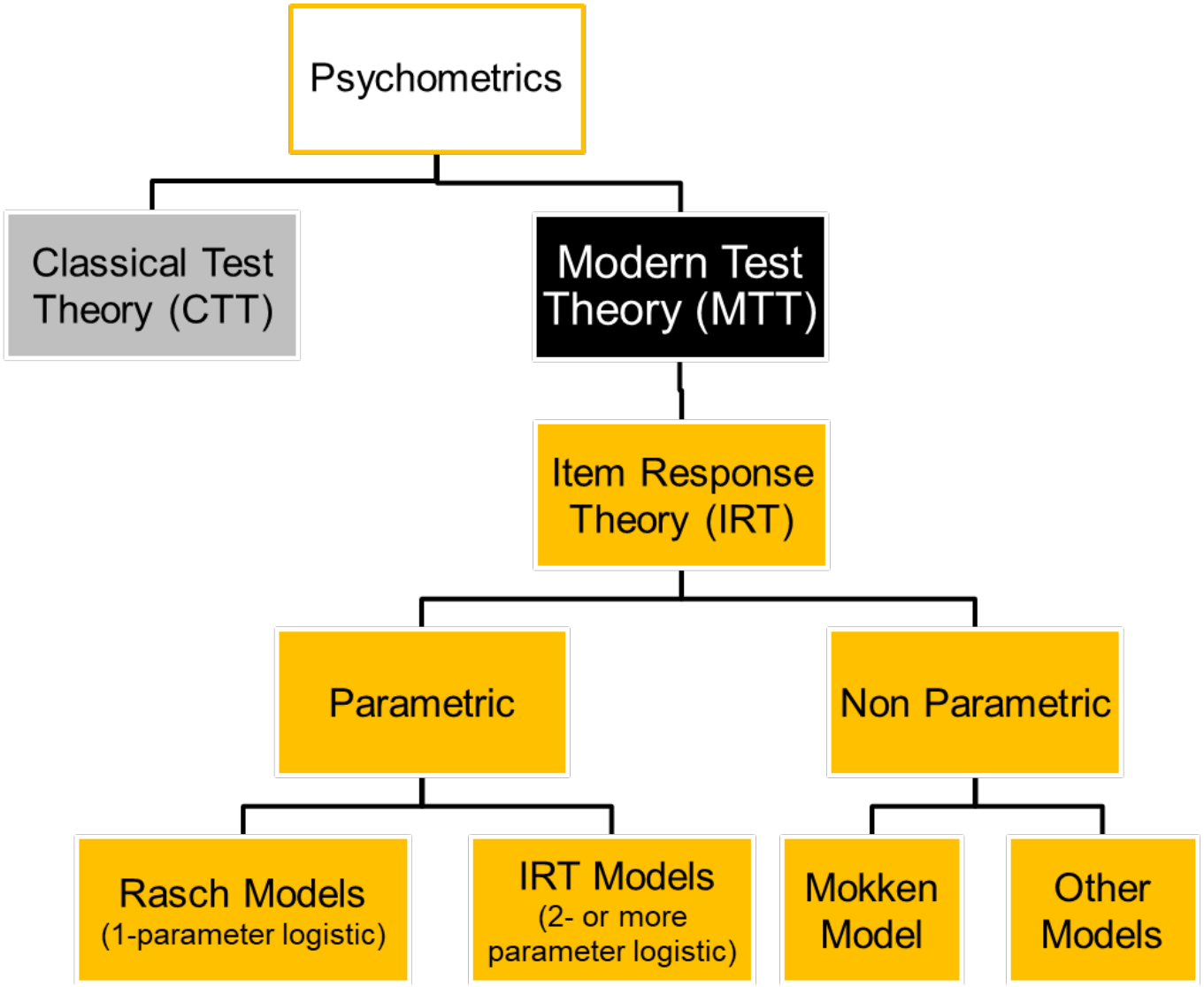
- This statistical method is named after the Danish Mathematician George Rasch who developed it[2]
- Statistical method from the field of modern/probabilistic test theory.
- Rasch analysis is mainly used to determine metric properties of questionnaires.
- It is widely used and especially in the fields of educational, psychological, and health assessment.
- Once fit to the Rasch model is established a score is considered interval scaled and a sufficient statistic for measurement and parametric analysis.

Rasch Analysis

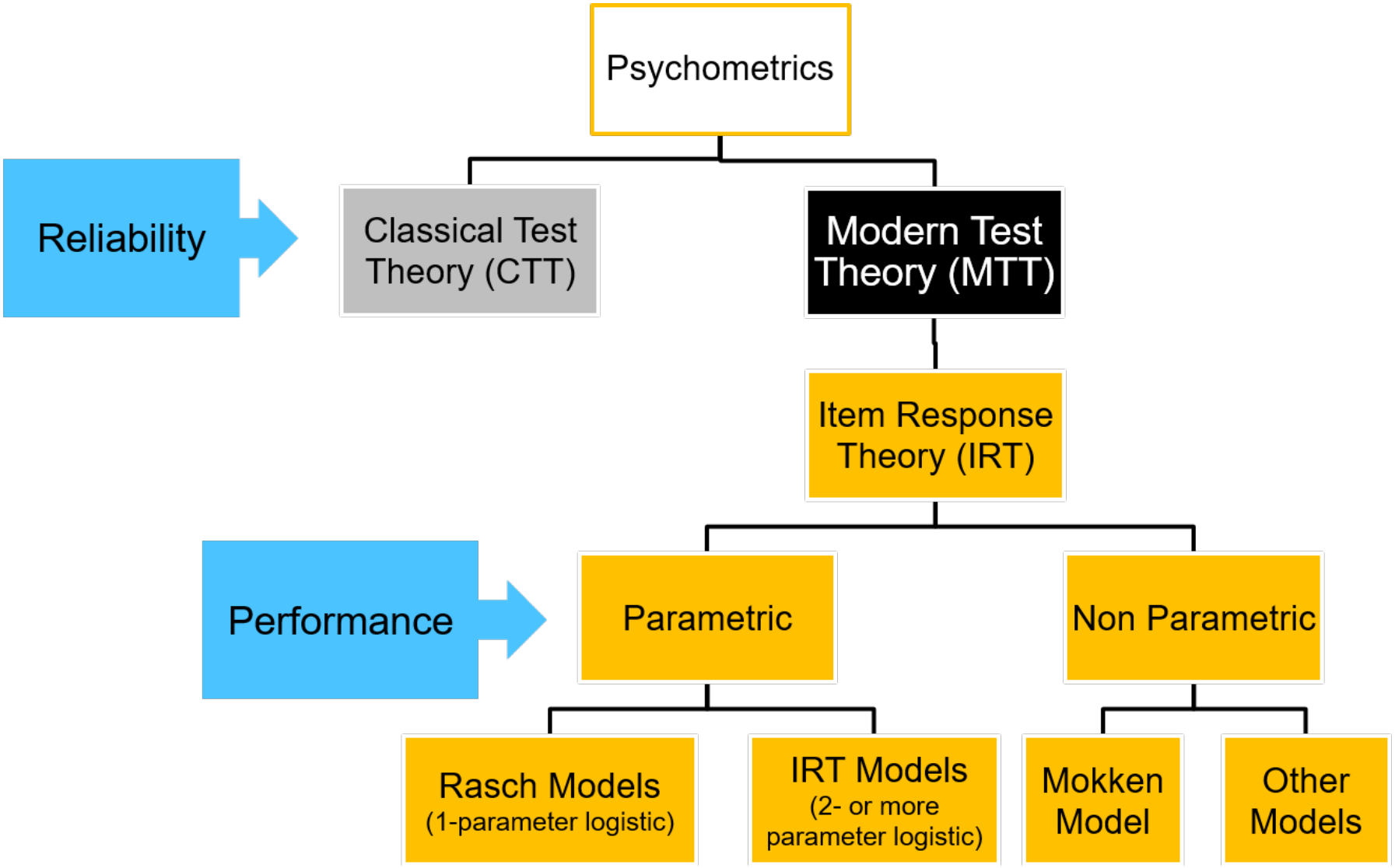
Rasch analysis is frequently applied to verify if an instrument shows important (psycho)metric properties.

- Stochastic ordering (fit of data to model)
- Monotonicity (ordering or response options)
- No local response dependencies or LID (no significant correlations between items)
- Unidimensionality (one latent construct)
- No differential item functioning or DIF (no sample subgroup effects)

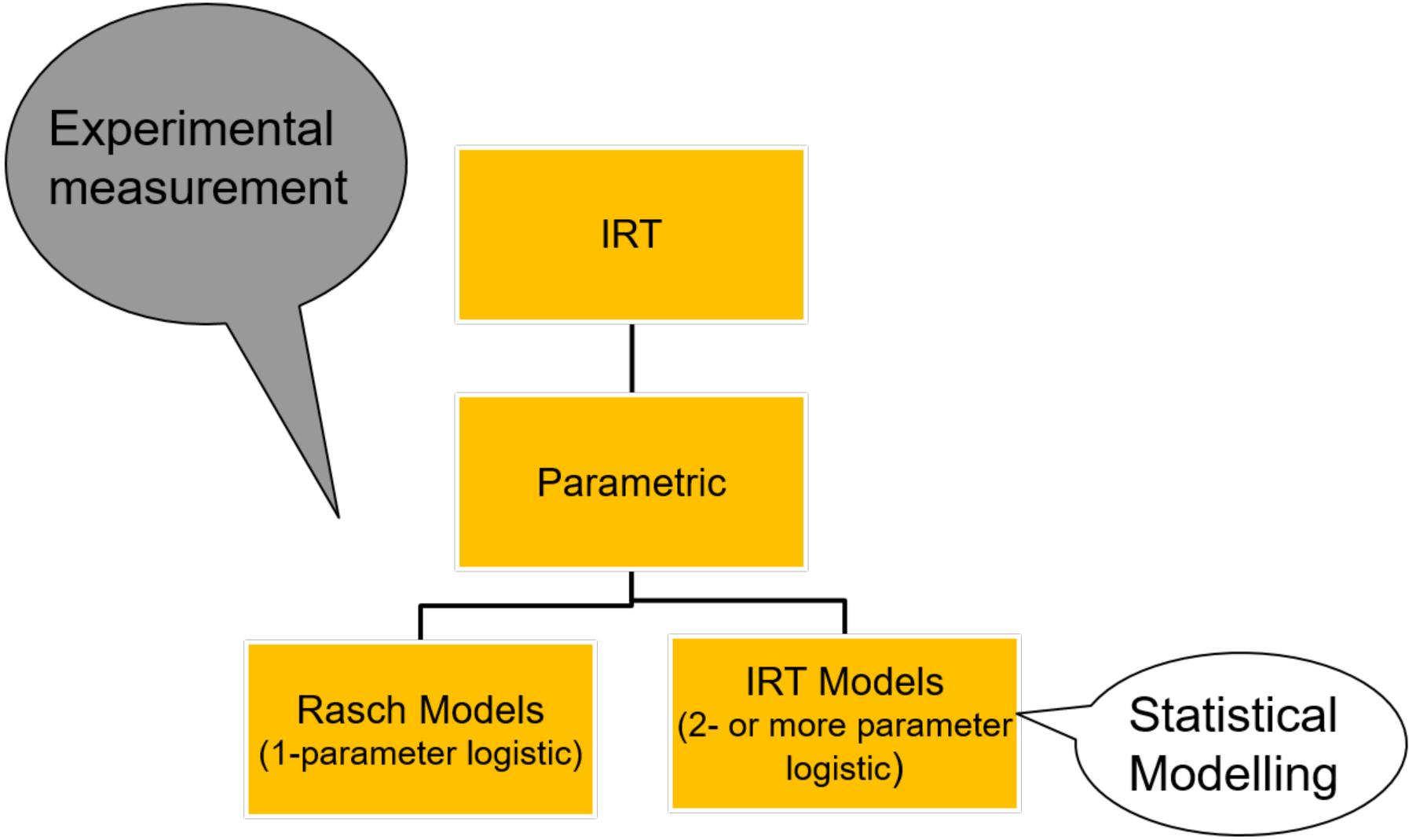
Psychometrics



Psychometrics

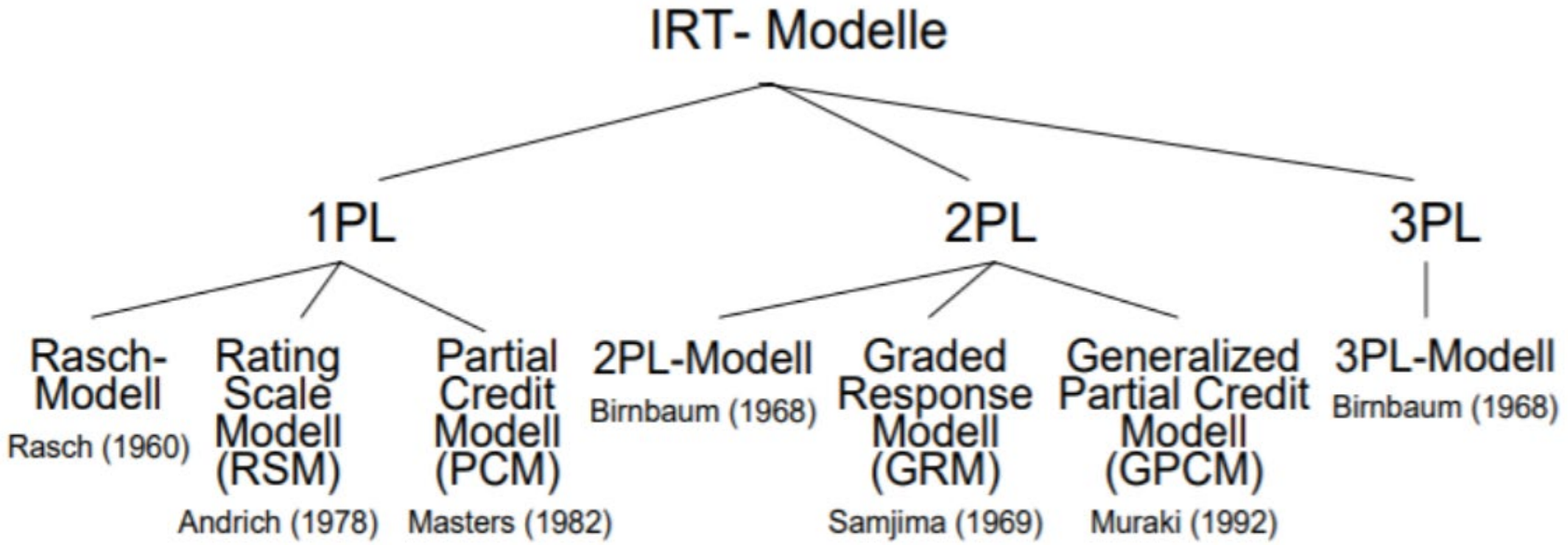


Two Competing Paradigms



Rasch vs IRT[1]

Family of Rasch and IRT Models[1]



[1] A simplified view

The Rasch Function

Rasch Model for dichotomous responses (Rasch 1960).

The probability of a person *j* with an ability θ to respond correctely (Xij = 1) to the item *i*, having difficulty β is formalized:

$$P(X_{ij} = 1|\theta_j) = \frac{exp(\theta_j - \beta_i)}{1 + exp(\theta_j - \beta_i)}$$

Rasch vs. IRT Equations

Rasch Equation

$$P(X_{ij} = 1|\theta_j) = \frac{exp(\theta_j - \beta_i)}{1 + exp(\theta_j - \beta_i)}$$

2-Parameter Logistic

$$P(X_{ij} = 1|\theta_j) = \frac{exp[a_i(\theta_j - \beta_i)]}{1 + exp[a_i(\theta_j - \beta_i)]}$$

Rasch vs. IRT Equations

Rasch Equation

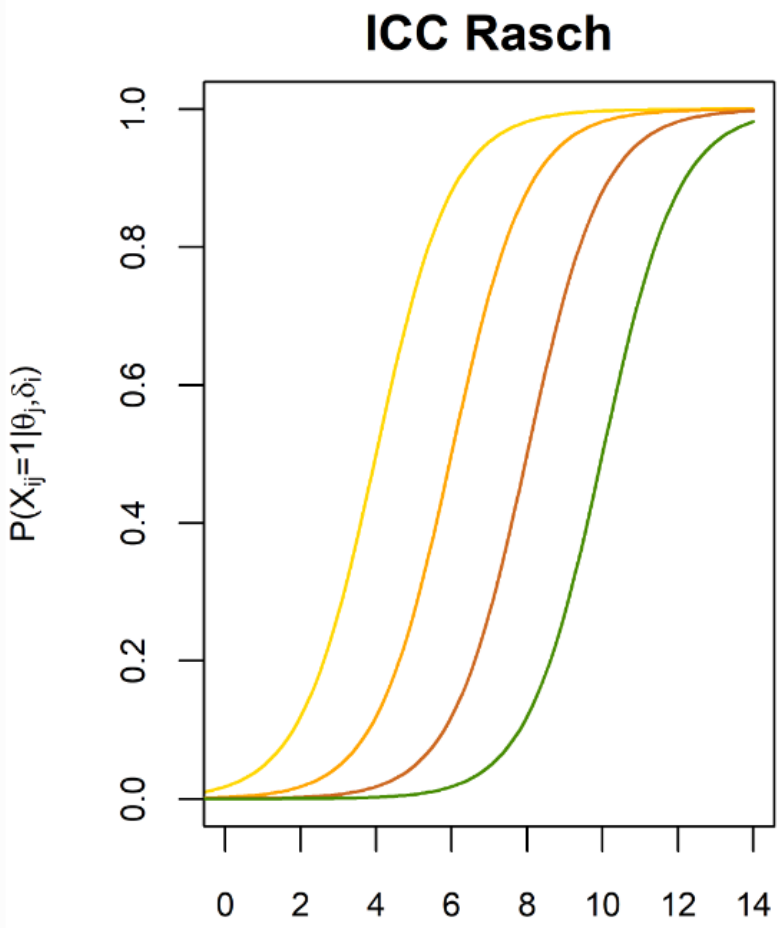
$$P(X_{ij} = 1|\theta_j) = \frac{exp(\theta_j - \beta_i)}{1 + exp(\theta_j - \beta_i)}$$

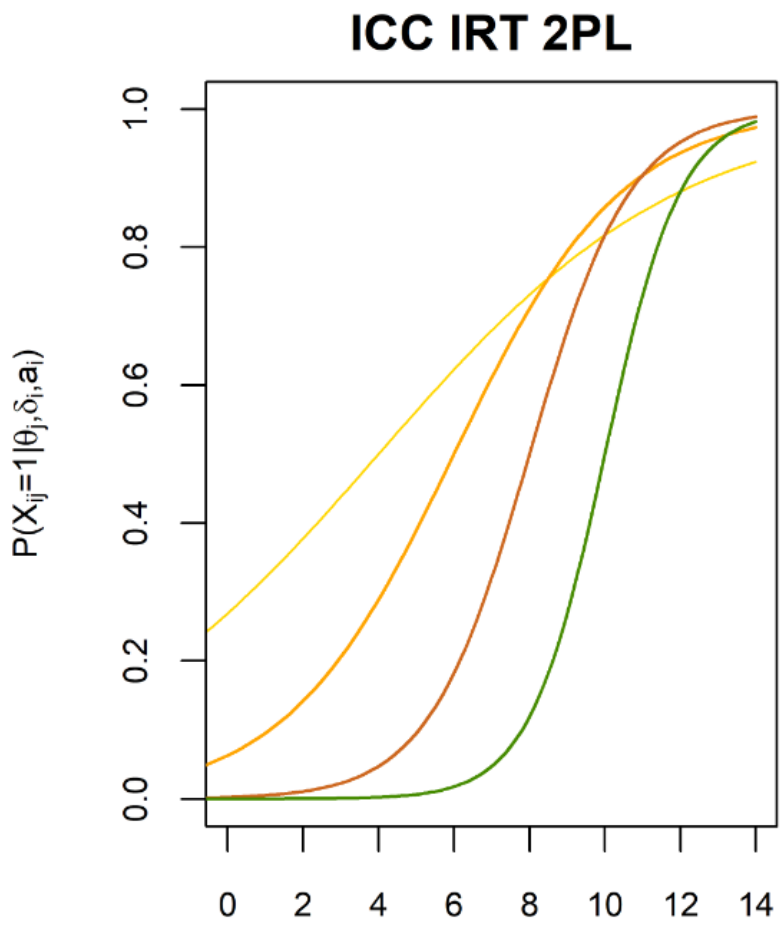
2-Parameter Logistic

$$P(X_{ij} = 1|\theta_j) = \frac{exp[a_i(\theta_j - \beta_i)]}{1 + exp[a_i(\theta_j - \beta_i)]}$$

a_i slope/discrimination parameter

Family of Rasch and IRT Models





Probabilistic Models of Measurement

Free parameter in the various models
Rasch Models : item difficulty
1-Parameter Logistic (1-PL) : item difficulty (very similar to the Rasch model)
2-Parameter Logistic (2-PL): item difficulty, item discrimination
3 Parameter Logistic (3-PL): item difficulty, item discrimination, guessing parameter
4 Parameter Logistic (4-PL): item difficulty, item discrimination, guessing parameter, ‘slipping parameter’.

Score Sufficiency

Rasch perspective: The raw score has all the information about the «ability» of the respondent. = **Score Sufficiency**

IRT perspective: The pattern of responses has all the information about the «ability» of the respondent.

Guttman Pattern

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	Score Item
J1											10
J2										x	9
J3									x	x	8
J4								x	x	x	7
J5							x	x	x	x	6
J6						x	x	x	x	x	5
J7					x	x	x	x	x	x	4
J8				x	x	x	x	x	x	x	3
J9			x	x	x	x	x	x	x	x	2
J10		x	x	x	x	x	x	x	x	x	1
Score Person	10	9	8	7	6	5	4	3	2	1	

High ability

Low ability

1+1= Easy Items

Difficult Item

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx =$$

Translating: Educational Measurement to Health Assessment with MDS

Rasch: The probability of a **correct response** is a function of a person's ability and an item's difficulty.

In Education:

Higher Ratings/More points on items (less mistakes) -> Higher Total Score -> More Ability.

In the MDS:
Capacity problems are rated 1 (None) and 5 (Extreme)
Performance difficulties are rated 1 (None) and 5 (Extreme)

In MDS (health Measurement): Higher Ratings on Items -> Higher Total Score -> More Ability?

Higher Ratings/More points -> Higher scores -> More problems

Guttman Pattern

Lot of points = low functioning

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	Score Item
J1											10
J2										x	9
J3									x	x	8
J4								x	x	x	7
J5							x	x	x	x	6
J6						x	x	x	x	x	5
J7					x	x	x	x	x	x	4
J8				x	x	x	x	x	x	x	3
J9			x	x	x	x	x	x	x	x	2
J10		x	x	x	x	x	x	x	x	x	1
Score Person	10	9	8	7	6	5	4	3	2	1	

High ability

Perfect! Almost 0 points!

Easy Items

Seeing without glasses

Difficult Item

Washing and dressing oneself

Summary

Measurement : different scale levels

Psychometrics:
Various models allow to calibrate scales and to obtain interval-scales properties.

Rasch analysis possess score sufficiency property: one score one ability. Universality of the metric.

Interpretation Rasch: think it twice -> what means more

Coming next:

- Assumptions of the Rasch Model
- Resources for Rasch analysis
- Rasch Analysis with WHOMDS
- Disaggregate Reporting in MDS

Summary

Part 2: Monday 19.04.2021 – MDS metrics

9.00 – 9.30 Welcome, Q&A for Part I

9:30 – 11:00 Assumptions of the Rasch Model

11:00 – 11:15 *15 Minutes: Questions and Coffee Break*

11:15 – 11:45 Rasch analysis – Resources: Software, Packages, Literature

R and R-Studio

Brief Introduction to R

11:45 – 12:45 Psychometric Analysis using the WHOMDS-package

12:45 – 13:00 *15 Minutes: Questions and End*

THANK YOUR FOR YOUR ATTENTION!