

$Korr(e_v, t_v) = 0$   
 $P_{2PL} = \frac{\exp(\beta(\theta - \sigma))}{1 + \exp(\beta(\theta - \sigma))}$      $Erw(\sum_v e_v) = 0$      $P = \frac{\exp(\theta - \sigma)}{1 + \exp(\theta - \sigma)}$      $x_v = t_v + e_v$      $P_{3PL} = \gamma + (1 - \gamma) \frac{\exp(\beta(\theta - \sigma))}{1 + \exp(\beta(\theta - \sigma))}$      $t_v = Erw(x_v)$

## Item-Response-Theorie: Rasch-Modell

### Zum Vergleich: Die Klassische Testtheorie (KTT)

#### Idee und Messprinzip

$$x_v = t_v + e_v$$

$x_v$  Testergebnis  
 $t_v$  „Wahrer Wert“  
 $e_v$  Messfehler

z.B. Punktzahl  
z.B. Intelligenz  
z.B. Müdigkeit

#### Voraussetzungen

- Messfehler muss vollkommen zufällig sein, d.h. nicht mit anderen Größen korrelieren.

Dann gilt:  $t_v = Erw(x_v)$

#### Probleme

- Häufig keine zufällige Messfehler (Korrelation)
- Kein Umgang mit fehlenden Werten möglich
- Keine verallgemeinerbare Personenfähigkeit
- Keine intervallskalierten Testwerte

### Das Rasch-Modell (1PL) als Messmodell der Item-Response-Theorie

#### Idee und Messprinzip

Manifestes Verhalten (Antwort) wird durch latente Variablen bestimmt:

$$\log \frac{P}{1-P} = \theta - \sigma$$

$P$  Wahrscheinlichkeit für richtige Lösung  
 $\theta$  Fähigkeit des Probanden  
 $\sigma$  Schwierigkeit der Aufgabe

#### Voraussetzungen

- Lösungswahrscheinlichkeit muss sich in guter Näherung verhalten wie

$$P = \frac{\exp(\theta - \sigma)}{1 + \exp(\theta - \sigma)}$$

- Vernachlässigbare Ratewahrscheinlichkeit
- Gleiche Trennschärfe für alle Aufgaben
- Große Varianz bei Probanden und Aufgaben

#### Probleme

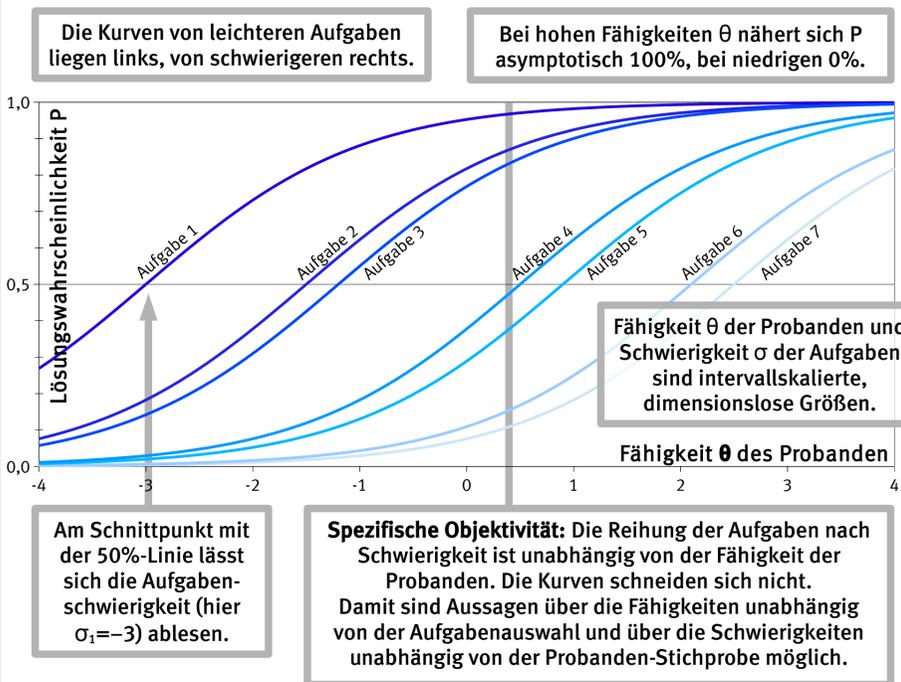
- Hohe Anforderungen an Aufgabenqualität
- Keine globalen Testgütekriterien

#### Eigenschaften des Modells

- Spezifische Objektivität (s.u. links)
- Aufgaben und Personen auf einer Skala
- Einfacher Umgang mit fehlenden Werten

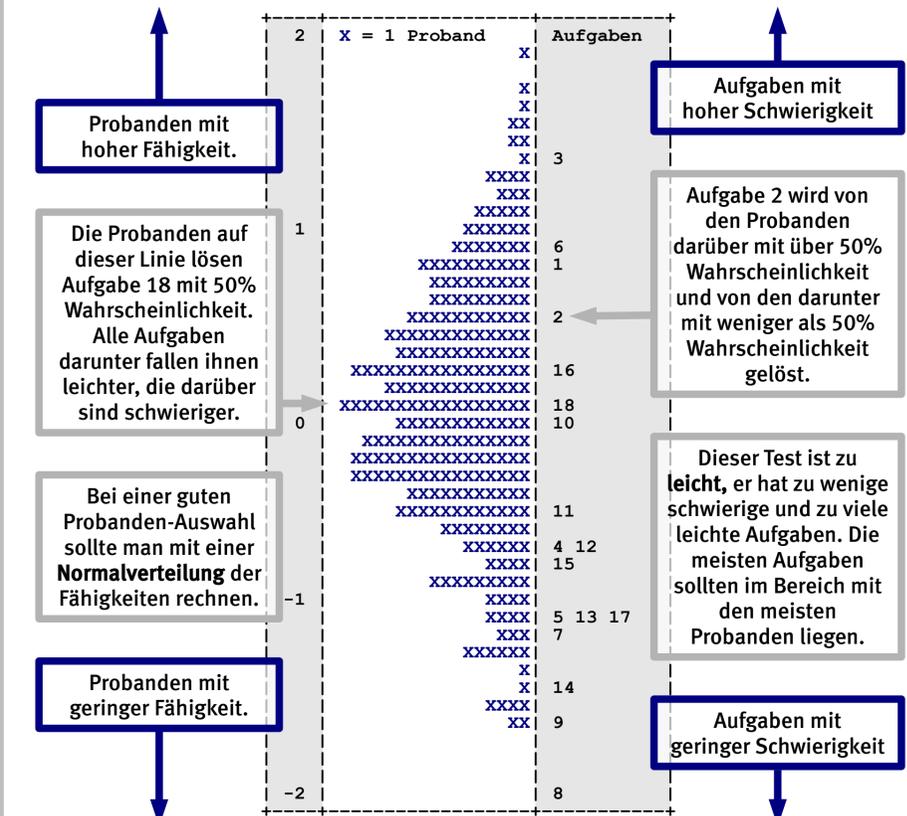
### Itemcharakteristik

In dieser Darstellung wird die Lösungswahrscheinlichkeit als Funktion der Fähigkeit des Probanden aufgetragen. (Formel im Kasten oben.)



### Wright-Map

Die Wright-Map stellt Probanden und Aufgaben einander gegenüber.



### Weitere IRT-Modelle

Die sog. Birnbaum-Modelle führen weitere Aufgaben-Parameter ein:

#### Zweiparametrisch-Logistisch (2PL)

Zusätzlich: Aufgaben-Diskrimination  $\beta$

$$P_{2PL} = \frac{\exp(\beta(\theta - \sigma))}{1 + \exp(\beta(\theta - \sigma))}$$

#### Dreiparametrisch-Logistisch (3PL)

Zusätzlich: Ratewahrscheinlichkeit  $\gamma$

$$P_{3PL} = \gamma + (1 - \gamma) \frac{\exp(\beta(\theta - \sigma))}{1 + \exp(\beta(\theta - \sigma))}$$

**Vorteil:** Bessere Passung des Modells auf die gegebenen Daten

**Nachteil:** Keine spezifischen Objektivität (Interpretation schwieriger)

### Ausblick

- Mehrdimensionalität durch mehrere latente Variablen pro Proband
- Komplexe Modelle: Aufgaben laden auf mehrere Dimensionen
- Mehrstufige Codierung von Aufgaben mit Partial Credit Modellen
- Graded Response Modelle ermöglichen likert-ähnliche Skalen
- Modellprüfung mit  $\chi^2$ -Test und BIC-, AIC-Vergleichen
- Prüfung der Modellkonformität von Aufgaben und Personen