

# **HEUREKO-CAT: Computergestütztes adaptives Assessment von Schülerkompetenzen beim heuristischen Arbeiten mit Repräsentationen funktionaler Zusammenhänge**

---



Markus Wirtz\*    Jan Henning-Kahmann\*

Timo Leuders<sup>Δ</sup>    Dominik Naccarella<sup>Δ</sup>

\***Institut für Psychologie**

<sup>Δ</sup>**Institut für Mathematische Bildung**



Bildungsforschung und Bildungsinformation

Ulf Kröhne (Technology Based Assessment)



Regina Bruder

**Fachbereich Mathematik**

# Hintergrund

## Gezielte Unterstützung der Diagnostik von Kompetenzen im Unterricht

- Strukturelle Entwicklung von Leistungsverläufen
- Identifikation von Kompetenzprofilen
- Identifikation fachdidaktisch bedeutsamer Teilkompetenzen

## Ziele von HEUREKO:


1. Konstruktion und Überprüfung eines Kompetenzstrukturmodells für mathematisches *Problemlösen* und *Modellieren* in Sekundarstufe I
2. Bereitstellung eines empirisch fundierten Diagnose- und Förderinstrumentariums für die Schulpraxis

# Inhaltsbereich

- Mathematische Beschreibung von „Wachstum und Veränderung“  
→ *Funktionales Denken*
- Kompetenzentwicklung im Inhaltsbereichs „Wachstum und Veränderung“ ist zentrales mathematisches Bildungsziel ab Klasse 7
- Zentral: heuristische Verwendung fundamentaler mathematischer **Darstellungsarten:**  
*numerisch,*  
*graphisch,*  
*symbolisch,*  
*situativ/verbal,* ... und der Wechsel zwischen ihnen!

# Beispiele für Aufgaben

**sonder ANGEBOT**

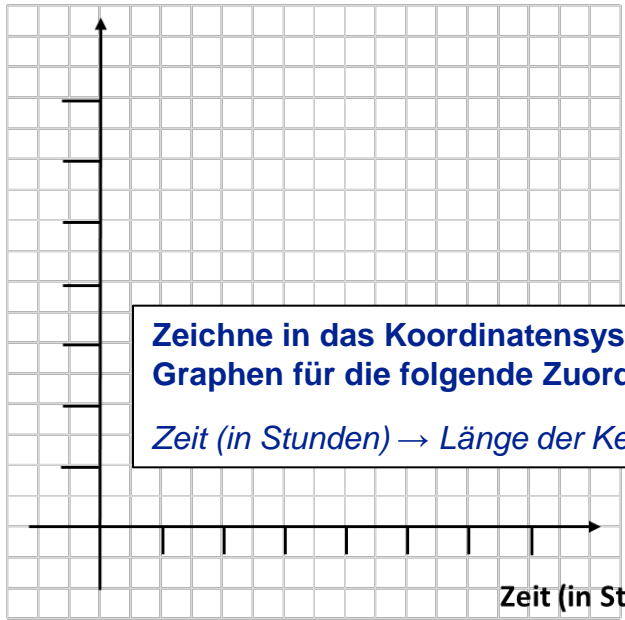


Extrem lange Brenndauer!  
 15 cm Kerze, 9 Stunden

Auf einem Joghurtbecher ist folgende Nährwerttabelle abgebildet:

**Vervollständige die Tabelle mit Hilfe der Angaben auf dem Joghurtbecher.**

Länge der Kerze (in cm)



**Zeichne in das Koordinatensystem einen Graphen für die folgende Zuordnung:**  
*Zeit (in Stunden) → Länge der Kerze (in cm)*

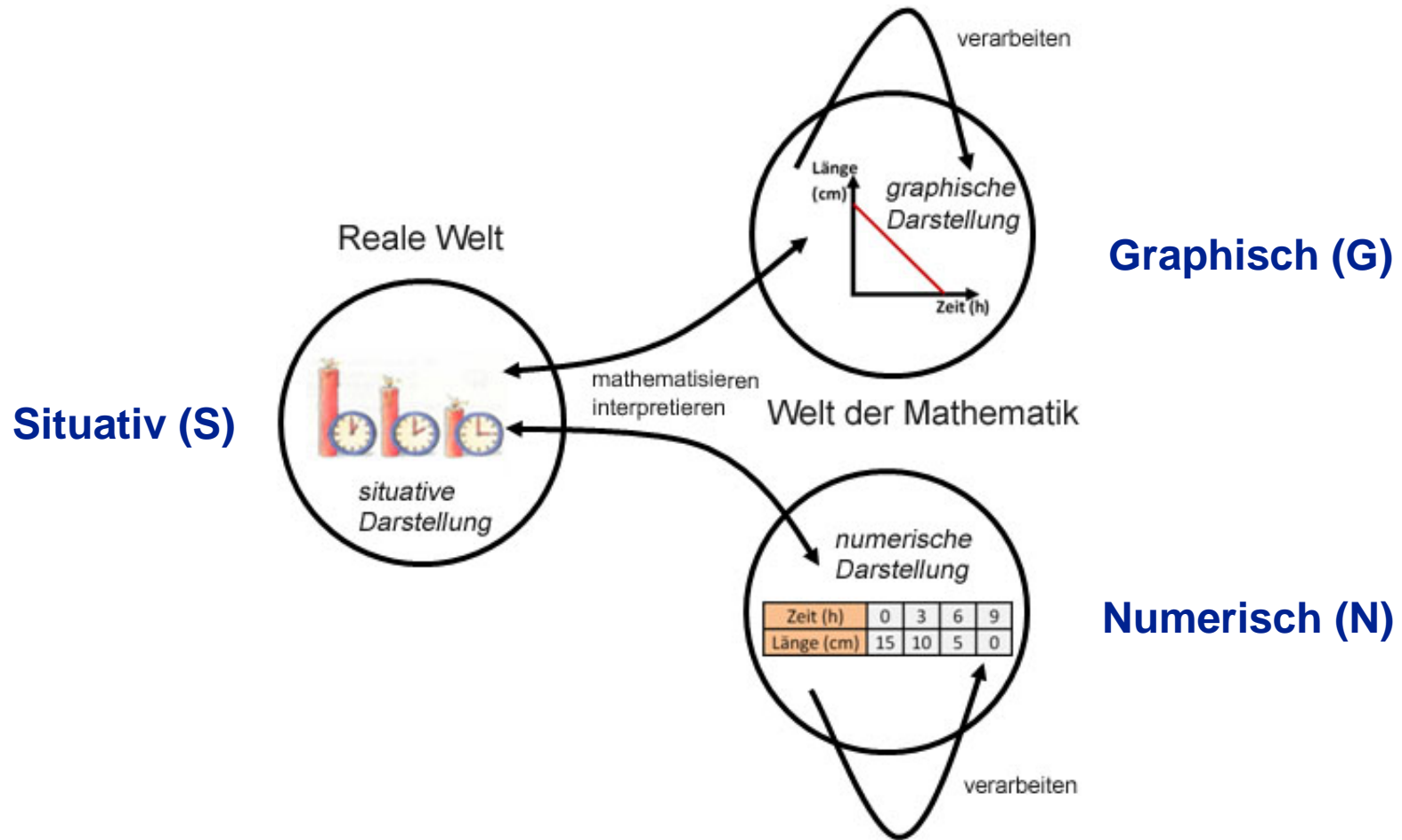
Durchschnittliche Nährwerte			
	Pro 100 ml	**Pro Portion (250 ml)	% GDA* (250 ml)
Brennwert	277 kJ 66 kcal	693 kJ 165 kcal	8 %
Eiweiß	3,4 g		
Kohlenhydrate	4,8 g		
Beim Zucker	4,8 g		
Fett	1,7 g		
Beim gesättigte Fettsäuren	2,3 g		
Natrium	0,8 g		

Joghurtmenge (in ml)	Brennwert (in kcal)	GDA in %
50		
100		
200		
250		

# Zentrale Forschungsfragen

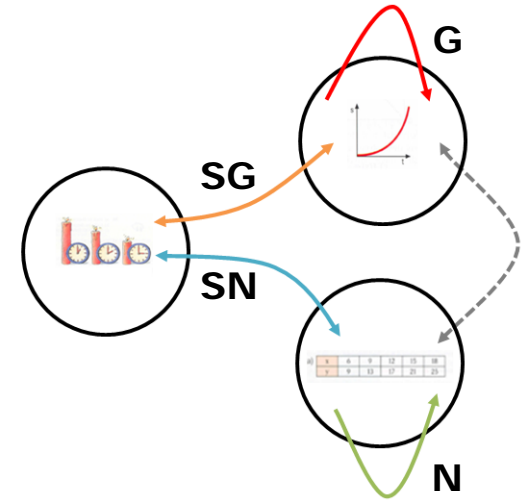
1. Welches **Kompetenzmodell** des Problemlösens und Modellierens im Bereich „Wachstum und Veränderung“ lässt sich aus bereits vorliegenden theoretischen und empirischen Erkenntnissen gewinnen?
2. Wie lassen sich die identifizierten Strukturkomponenten **psychometrisch** fundiert und differenziert **erfassen**?
3. Wie kann das entwickelte **Messinstrument** – v.a. für den Einsatz in der Schulpraxis – möglichst **effektiv** und **effizient** gestaltet werden?

# Kompetenzstruktur



# Problemstellung

- ✓ **Phase I + II:**  
Vierdimensionales Kompetenzmodell auf Basis eines curricular validen Itempools (103 Items)  
1. **S**ituativ-**G**raphisch, 2. **S**ituativ-**N**umerisch, 3. **G**raphisch, 4. **N**umerisch
- ✗ Reliabilität der Skalen für **Individualdiagnostik** noch nicht ausreichend hoch!
- ✗ **Testökonomie** (Zeitaufwand für Vorbereitung, Bearbeitung, Auswertung, Rückmeldung) noch optimierbar!
- ✗ Testpersonen sollten nur jene Aufgaben bekommen, die für sie den größten Informationsgewinn bringen!

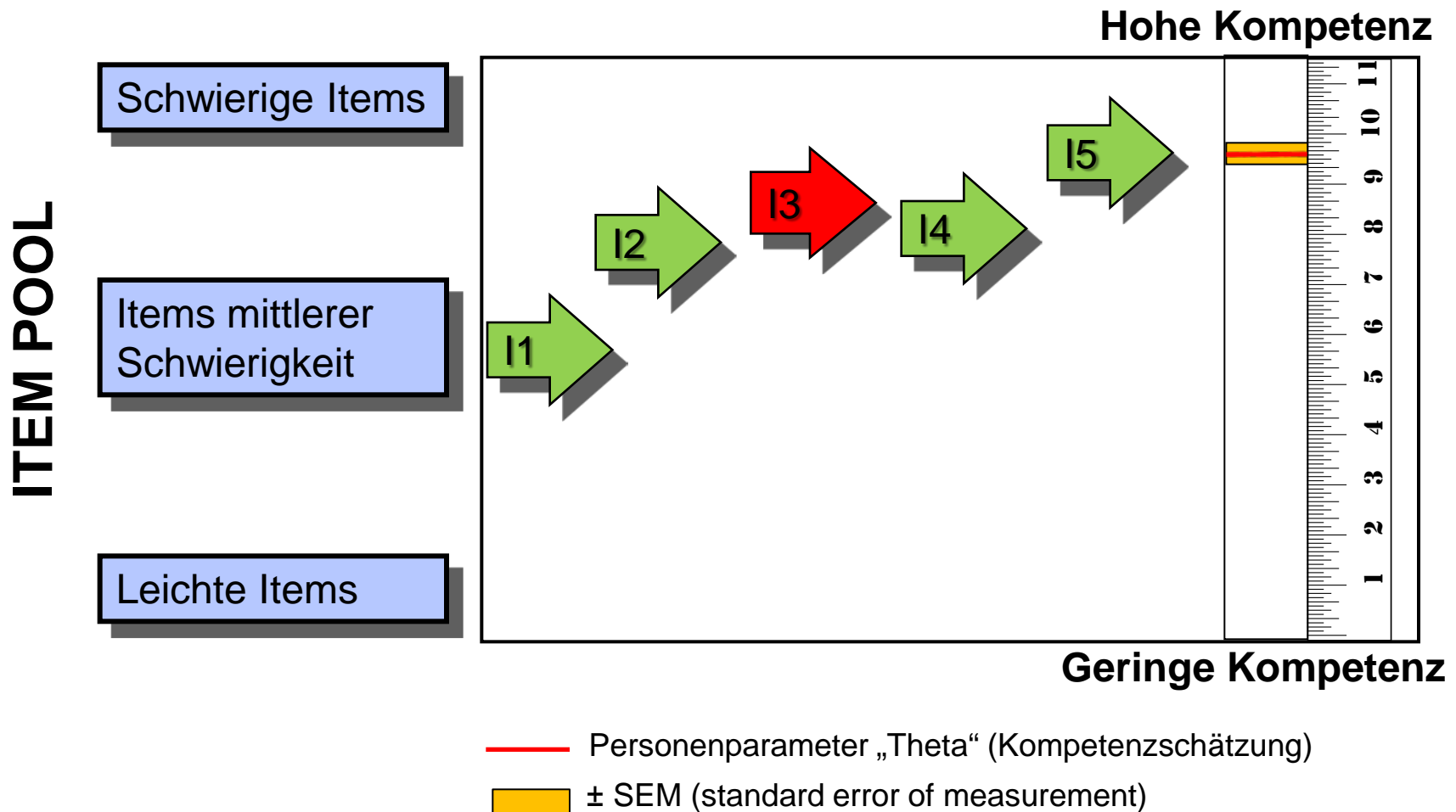


# Adaptives Testen

- Auswahl der zu bearbeitenden Testitems ist **abhängig** vom vorhergehenden Antwortverhalten der Testperson
- **Iterative Prozedur**: Fähigkeit der Testperson wird nach jedem beantworteten Item erneut geschätzt, d.h. „aktualisiert“
- **Auswahlkriterium**: maximale Präzisionszunahme der Fähigkeits-schätzung der Testperson (= hohe Messeffizienz)!
- Deutliche **Verkürzung der Testdauer** (Ökonomie) bei gleichzeitig gut definierter **Schätzgenauigkeit** (Reliabilität) möglich!



# Adaptives Testen



# Voraussetzungen adaptiven Testens

1. Items müssen **kalibriertem Itempool** entstammen, der Kriterien eines Modells der *Item Response Theory* (IRT) genügt
2. Ausreichende **Größe** des Itempools sowie **Abdeckung** des Kompetenzbereichs der Zielpopulation durch Itemschwierigkeiten
3. Hinreichende Itemdiskrimination (bei 2PL-, 3PL-Modell)
4. Items *sollten* als **computerbasierte Testaufgaben** administriert werden können (*computerized-adaptive testing* = CAT)

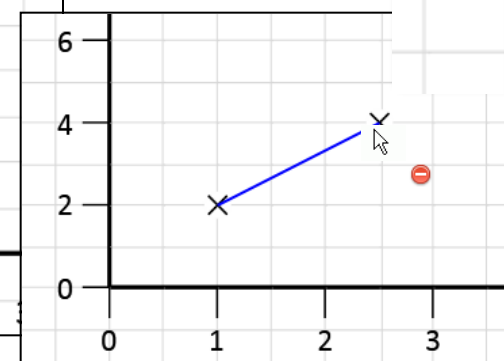
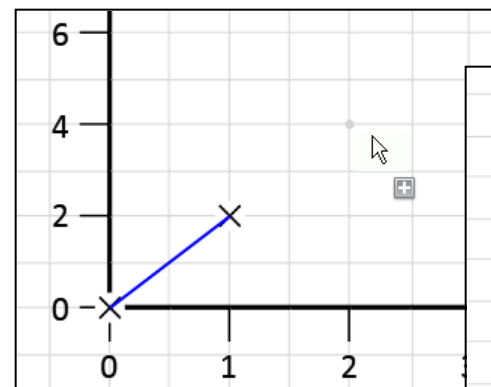
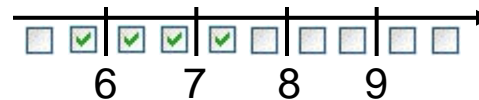
# Ziele

1. **Computerisierung** der papierbasierten Items in den Kompetenzdimensionen **Situativ-Graphisch** und **Situativ-Numerisch**
2. **Erweiterung** des Itempools auf  $n > 60$  Items pro Dimension, möglichst homogene **Abdeckung** des Schwierigkeitsspektrums
3. Empirische **Validierung und Skalierung** (IRT) des erweiterten computerisierten Itempools (*SG & SM*)
4. Entwicklung, Validierung und Implementierung einer **computerbasierten adaptiven Testprozedur** (CAT)

# Computerisierung der Items

## Umgesetzte Itemformate:

1. Multiple Choice
2. Bereichsmarkierung
3. „Labeling“ von Diagrammen, Tabellen
4. Graphen ergänzen
5. Zahleneingabe
6. Lückentext
7. Tabellen ergänzen
8. Tabellen erstellen
9. Graphen zeichnen



Bitte wählen ▾

Bitte wählen

- gefahrene Strecke
- Geschwindigkeit des Schiffes
- Laufzeit des Schalls
- Meerestiefe

100

200

250

300

▾

# Erweiterung des Itempools

**N= 166 Items** (SG: 91, SN: 75)

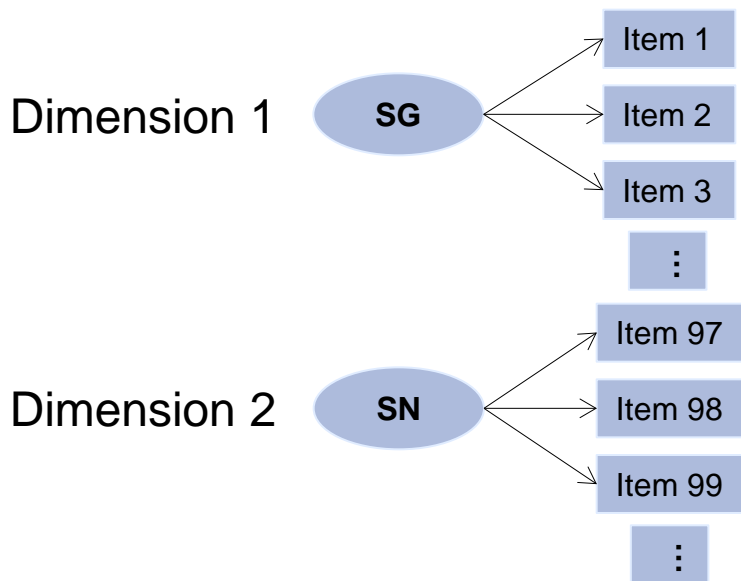
Aufgabencharakter

	lokal	vernetzt	offen
<b>(Kognitive) Aktivität</b>	<b>Ablesen</b> von Werten aus gegebenen - Graphen - Tabellen	<b>Erkennen / Labeln</b> einer Funktion in gegebenen - Graphen - Tabellen - Aussagen	<b>Ergänzen / Erstellen</b> von: - Graphen (Zeichnen) - Wertetabellen (Ausfüllen) für eine Funktion.
<b>leicht</b> z.B....	... einzelner, diskreter Wert ... Wertebereich	... geringe Anzahl an Distraktoren (2) ... einfache Funktion	... einfache lineare Funktion ... Werte fortsetzen
<b>mittel</b> z.B....	... interpolierter Wert ... mehrere, diskrete Werte	... mittlere Anzahl an Distraktoren (3-4) ... komplexere Funktion	... komplexere lineare Funktion ... Werte ergänzen
<b>schwer</b> z.B....	... extrapolierter Wert	... hohe Anzahl an Distraktoren (>4) ... komplexe Funktion	... komplexe lineare Funktion (exponential) ... neu erstellen

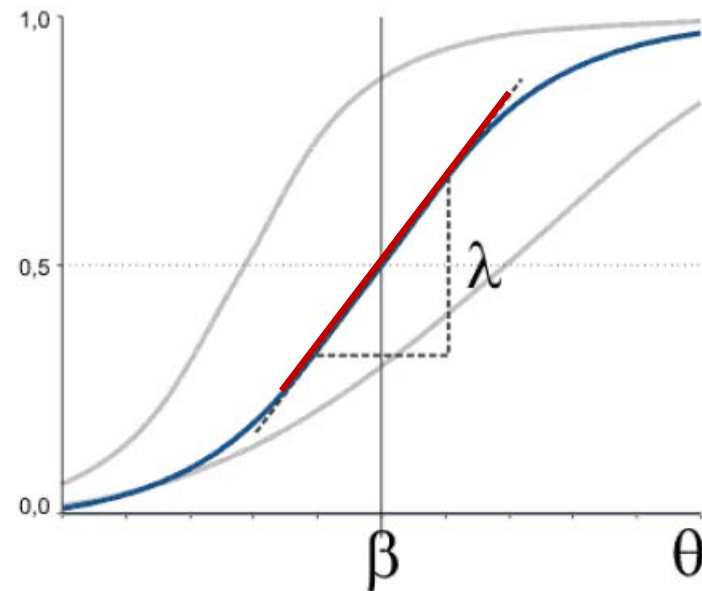
Komplexität

# IRT-Skalierung

**Modell-Fit:** zwei-dimensionales 2PL-Modell



Birnbaum-Modell (2PL)



# IRT-Skalierung

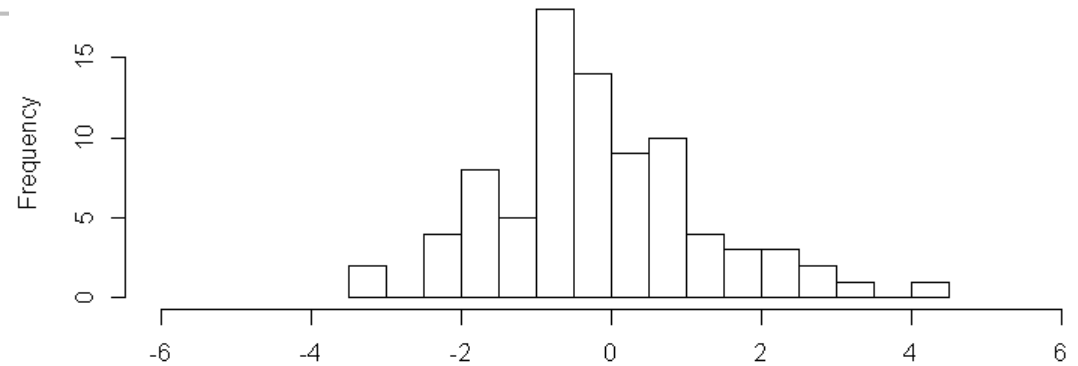
**Modell-Fit:** zwei-dimensionales 2PL-Modell

**Itemschwierigkeiten:**

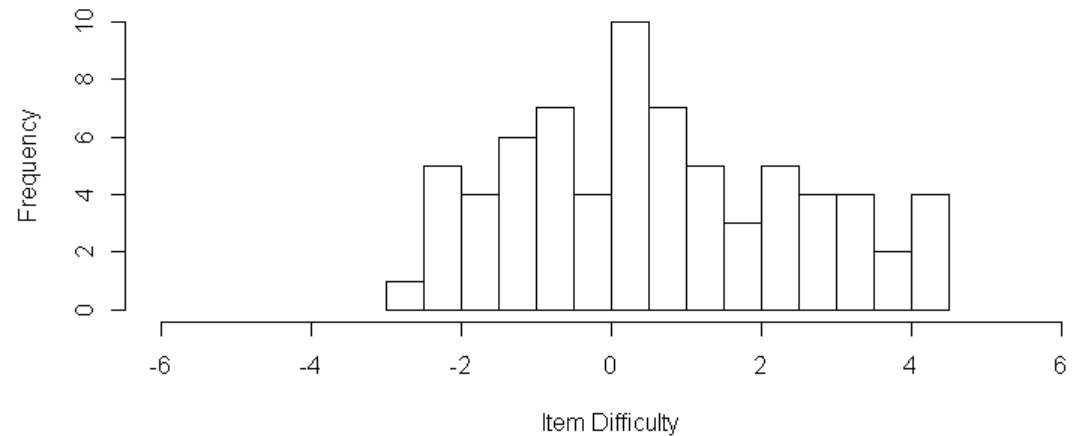
SG: -3.13 – 4.39

SN: -2.75 – 4.49

**Distribution of Item Difficulties (SG)**



**Distribution of Item Difficulties (SN)**



# IRT-Skalierung

**Modell-Fit:** zwei-  
dimensionales 2PL-Modell

**Itemschwierigkeiten:**

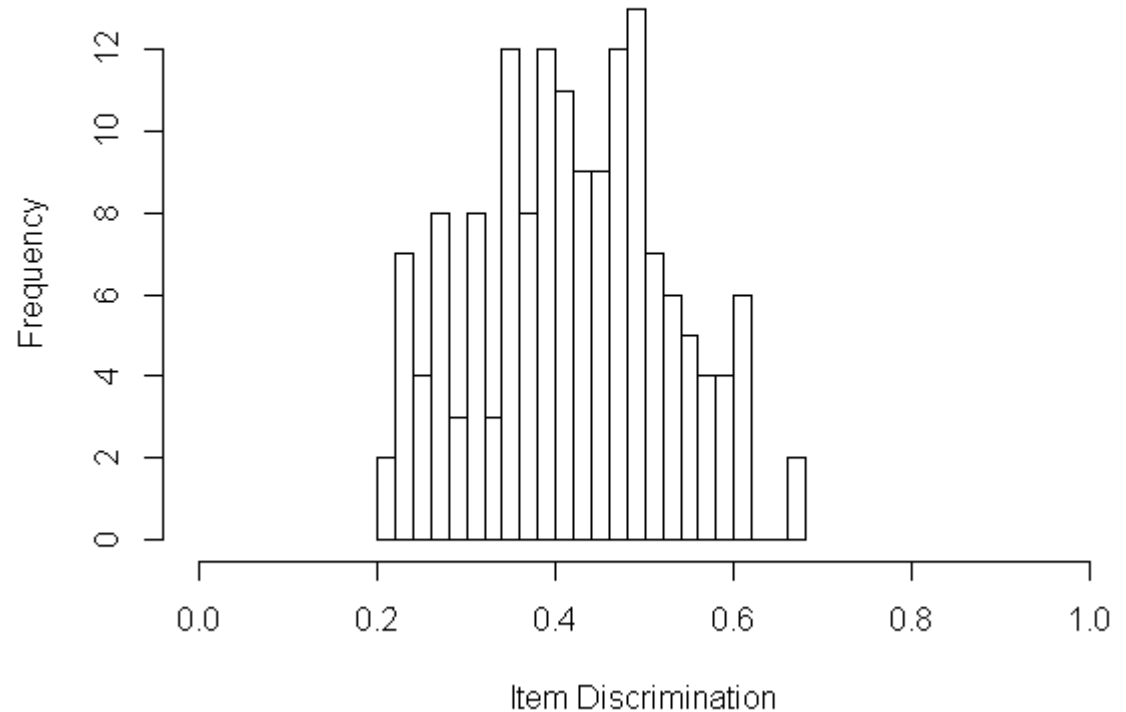
SG: -3.13 – 4.39

SN: -2.75 – 4.49

**Itemtrennschärfe (ktt):**

0.21 – 0.67

**Distribution of Item Discriminations in 2PL-Model**





# IRT-Skalierung

**Modell-Fit:** zwei-dimensionales 2PL-Modell

**Itemschwierigkeiten:**

SG: -3.13 – 4.39

SN: -2.75 – 4.49

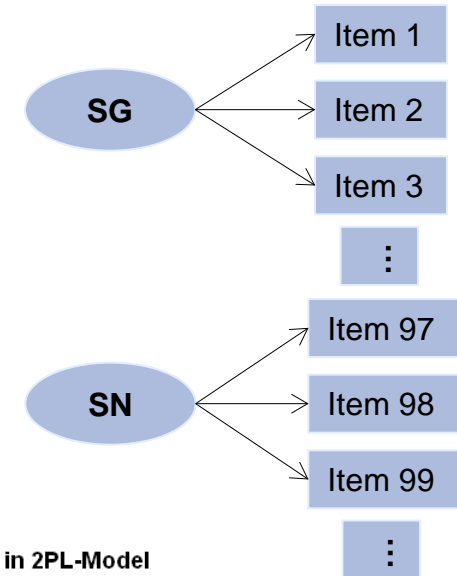
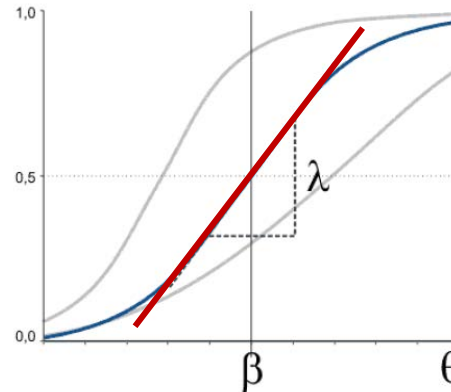
**Itemtrennschärfe (ktt):**

0.21 – 0.67

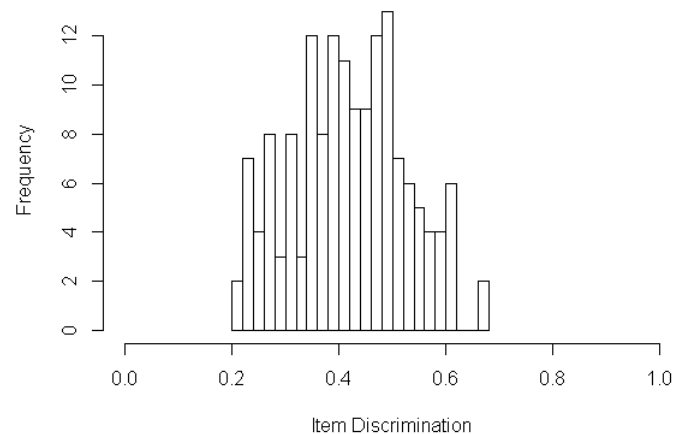
**EAP Reliabilität:**

SG: 0.78

SN: 0.78



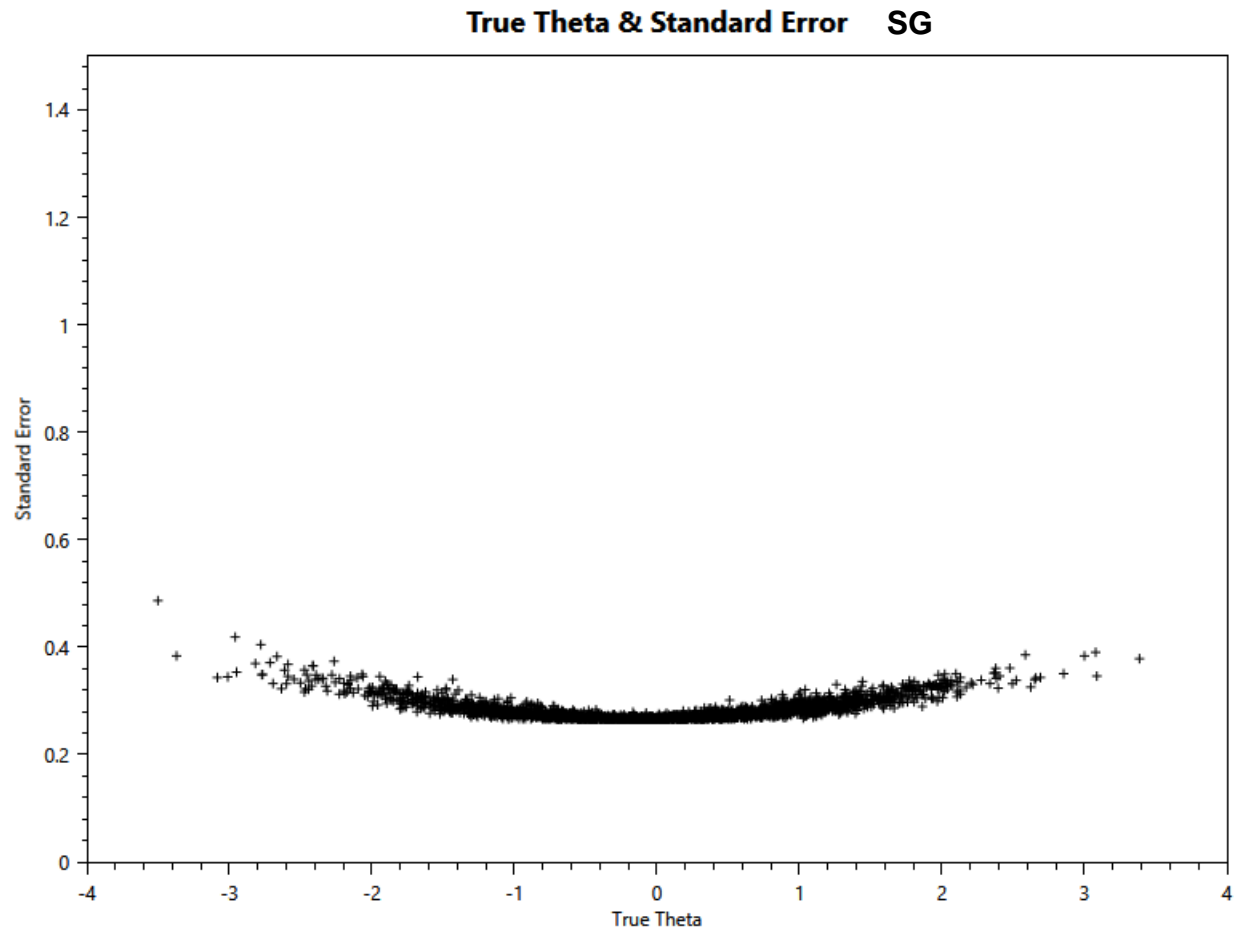
Distribution of Item Discriminations in 2PL-Model



# CAT-Simulation in MATE

## CAT-Optionen:

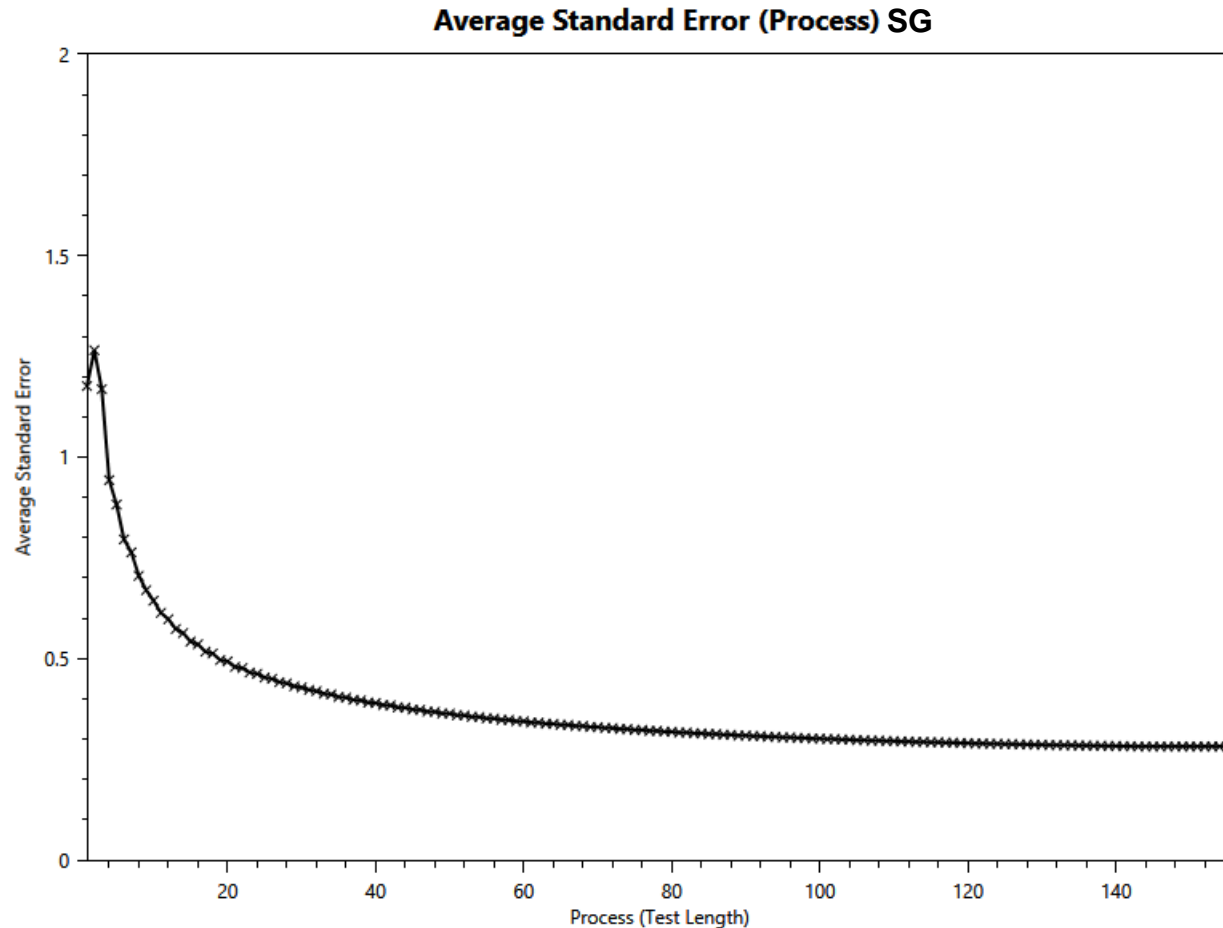
- MAT (multidimensional adaptive testing)
- Itemauswahl: maximum information
- „Content balancing“ (SG/SN)
- N= 1000
- abgebildet: SG



# CAT-Simulation in MATE

## CAT-Optionen:

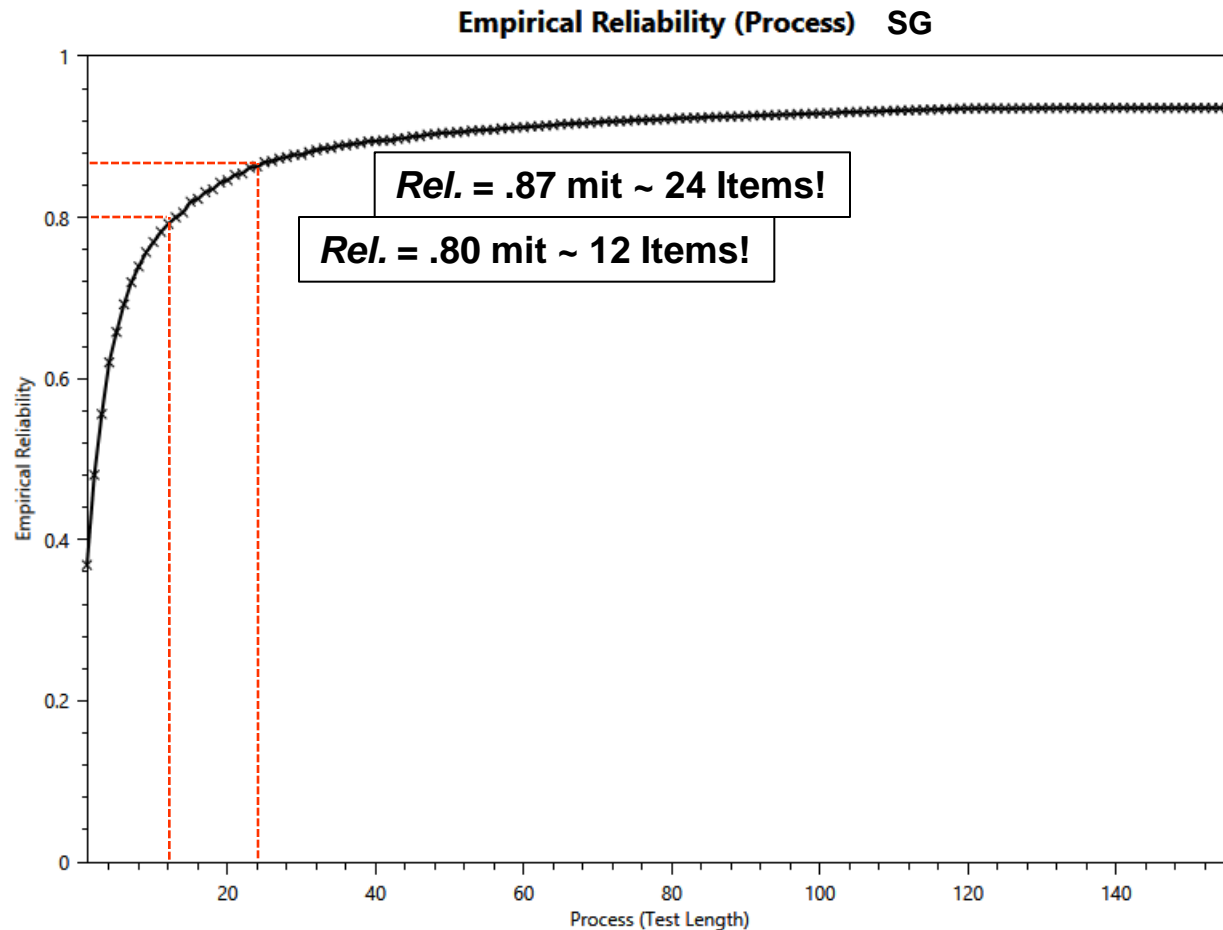
- MAT (multidimensional adaptive testing)
- Itemauswahl: maximum information
- „Content balancing“ (SG/SN)
- N= 1000
- abgebildet: SG



# CAT-Simulation in MATE

## CAT-Optionen:

- MAT (multidimensional adaptive testing)
- Itemauswahl: maximum information
- „Content balancing“ (SG/SN)
- N= 1000
- abgebildet: SG



## Fazit

- ✓ Vierdimensionales Kompetenzmodell (*SG, SN, G, M*) auf Basis eines curricular validen Itempools
- ✓ Hohe Testökonomie durch computer-based assesment (CBA)
- ✓ Hohe Reliabilität (v.a. für Individualdiagnostik) der Skalen *SG* & *SN* durch computer-adaptive Testprozedur (CAT)
- Anschlussfähigkeit an vorherige und Folgekompetenzen bleibt zu klären...
- Validierung des CAT in Schulpraxis wünschenswert!

# Kontakt:

**jan.henning@ph-freiburg.de**  
**markus.wirtz@ph-freiburg.de**

[www.ph-freiburg.de/psychologie/abteilungen/forschungsmethoden](http://www.ph-freiburg.de/psychologie/abteilungen/forschungsmethoden)