

**Möglichkeiten und Grenzen einer  
Kompetenzdiagnostik in der Soziologie:  
Ergebnisse einer Piloterhebung zur Messung  
von Kompetenzen in Methoden der empirischen  
Sozialforschung**

**Felix Wolter & Jürgen Schiener**

Institut für Soziologie  
Johannes Gutenberg-Universität Mainz

# Gliederung

1. **Motivation**
2. **Begriffe, Konzepte und Methoden der Kompetenzforschung**
3. **Studiendesign**
  - Kompetenzmodell
  - Implementation und Ablauf
4. **Ergebnisse**
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. **Fazit, offene Fragen und Ausblick**

## Projekt „Messung soziologischer Kompetenzen“

- **2012–2013: erste Pilotstudie („soziologische Basiskompetenzen“)**
  - Vortrag letztes Jahr
- **2013–2014: zweite Pilotstudie („Kompetenzen in (quantitativen) Methoden der empirischen Sozialforschung“)**
  - dieser Vortrag
- **Geplant: größerer Drittmittelantrag**

## Projekt „Messung soziologischer Kompetenzen“

- **2012–2013: erste Pilotstudie („soziologische Basiskompetenzen“)**
  - Vortrag letztes Jahr
- **2013–2014: zweite Pilotstudie („Kompetenzen in (quantitativen) Methoden der empirischen Sozialforschung“)**
  - dieser Vortrag
- **Geplant: größerer Drittmittelantrag**

# Gliederung

## 1. Motivation

## 2. Begriffe, Konzepte und Methoden der Kompetenzforschung

## 3. Studiendesign

- Kompetenzmodell
- Implementation und Ablauf

## 4. Ergebnisse

- Skalierung der Kompetenzindikatoren
- Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial

## 5. Fazit, offene Fragen und Ausblick

## Warum eine Kompetenzmessung in der Soziologie?

- **Hintergrund:**
  - Seit ~2000 Output-Orientierung („empirische Wende“) in der Bildungsevaluation
  - regelmäßige Kompetenzdiagnostik im schulischen Bereich  
(→ PISA et al.)
  - mittlerweile gut ausgearbeitete Kompetenzmodelle und Messinstrumente
- **Dagegen: Erheblicher Forschungsbedarf für den Hochschulbereich**
  - Aktuell: BMBF-Forschungsprogramm „Kompetenzen im Hochschulsektor“ mit >20 Projekten
  - Bisher kaum Forschung zu „soziologischen Kompetenzen“

## Warum eine Kompetenzmessung in der Soziologie?

- **Drei Motivationen für eine Kompetenzmessung in der Soziologie:**

## Warum eine Kompetenzmessung in der Soziologie?

- **Drei Motivationen für eine Kompetenzmessung in der Soziologie:**
  - Hochschulevaluation:
    - Aktuelle Evaluationspraxis (anhand von Input-Indikatoren und subjektiver Indikatoren) unzureichend oder zumindest problematisch.
    - Viele Curricula (Modulhandbücher) schreiben den Erwerb bestimmter Kompetenzen als Ziel fest, ohne dass bisher das Erreichen der Ziele objektiv gemessen oder überprüft werden kann.

## Warum eine Kompetenzmessung in der Soziologie?

- **Drei Motivationen für eine Kompetenzmessung in der Soziologie:**
  - Hochschulevaluation:
    - Aktuelle Evaluationspraxis (anhand von Input-Indikatoren und subjektiver Indikatoren) unzureichend oder zumindest problematisch.
    - Viele Curricula (Modulhandbücher) schreiben den Erwerb bestimmter Kompetenzen als Ziel fest, ohne dass bisher das Erreichen der Ziele objektiv gemessen oder überprüft werden kann.
  - Inhaltlich-soziologisches Argument:
    - „Kompetenz“ als besserer Indikator für (ungleichheits)soziologische Bildungsforschung (Determinanten und Erträge von Bildung).
    - Bisher: Problematische Proxy-Indikatoren wie Zertifikate oder „Noten“.

## Warum eine Kompetenzmessung in der Soziologie?

- **Drei Motivationen für eine Kompetenzmessung in der Soziologie:**
  - Hochschulevaluation:
    - Aktuelle Evaluationspraxis (anhand von Input-Indikatoren und subjektiver Indikatoren) unzureichend oder zumindest problematisch.
    - Viele Curricula (Modulhandbücher) schreiben den Erwerb bestimmter Kompetenzen als Ziel fest, ohne dass bisher das Erreichen der Ziele objektiv gemessen oder überprüft werden kann.
  - Inhaltlich-soziologisches Argument:
    - „Kompetenz“ als besserer Indikator für (ungleichheits)soziologische Bildungsforschung (Determinanten und Erträge von Bildung).
    - Bisher: Problematische Proxy-Indikatoren wie Zertifikate oder „Noten“.
  - Außendarstellung:
    - Berufsbild und Inhalt der soziologischen Ausbildung sind in der Öffentlichkeit und der Berufswelt immer noch vage und unklar. Daher: Wissen und Fähigkeiten von Soziolog/innen sprachlich fixieren und messbar machen.



# Gliederung

1. Motivation
2. **Begriffe, Konzepte und Methoden der Kompetenzforschung**
3. Studiendesign
  - Kompetenzmodell
  - Implementation und Ablauf
4. Ergebnisse
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. **Fazit, offene Fragen und Ausblick**

# Zum Kompetenzbegriff

- **Gängige Definition von „Kompetenzen“:**
  - „*kontextspezifische kognitive Leistungsdispositionen*, die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme/Leutner 2006: 879, Herv. i. Orig.).
  - Handeln können („Can-Do-Aussagen“)
  - Erlernbarkeit (im Unterschied zu Konstrukten wie „Intelligenz“)

# Kompetenzstrukturmodelle

- **Generell:**
  - beziehen sich auf die Kompetenzinhalte
  - Dimensionen und Subdimensionen
  - kognitive Prozesse und fachliche Gegenstände (Krathwohl 2002)
  - Festlegung von Kompetenzstrukturmodellen ist eine normative Entscheidung
- **Mögliche Vorgehensweisen / Quellen im Hochschulbereich:**
  - Analyse von Curricula, Modulhandbüchern, Lehrveranstaltungsplänen, Lehrbücher, Experteninterviews mit Fachvertretern, Festlegungen / Empfehlungen von Gremien, Verbänden, Beiräten usw.
  - Definition ausgehend von typischen Anforderungen auf dem Arbeitsmarkt

# Methoden: Item Response Theorie (IRT)

- **Idee und Ziele:**
  - Schluss von beobachtbaren Probandenantworten auf Testitems auf zugrundeliegende, nicht beobachtbare latente Variable (=Kompetenz).
  - Itemhomogenität (lokale stochastische Unabhängigkeit) bzw. Eindimensionalität: Nur das gewünschte latente Merkmal (und die Schwierigkeit der Aufgaben) beeinflusst die Antworten auf Testitems.
- **Zentrale Modelle:**
  - Rasch- bzw. 1PL-Modell (Rasch 1960)
  - Birnbaum- bzw. 2PL-Modell (Birnbaum 1968)
  - Rateparameter-Modell bzw. 3PL-Modell und erweiterte Modelle
  - Partial-Credit-Modelle für ordinale Items (Master 1982, Muraki 1992)

# Methoden: Item Response Theorie (IRT)

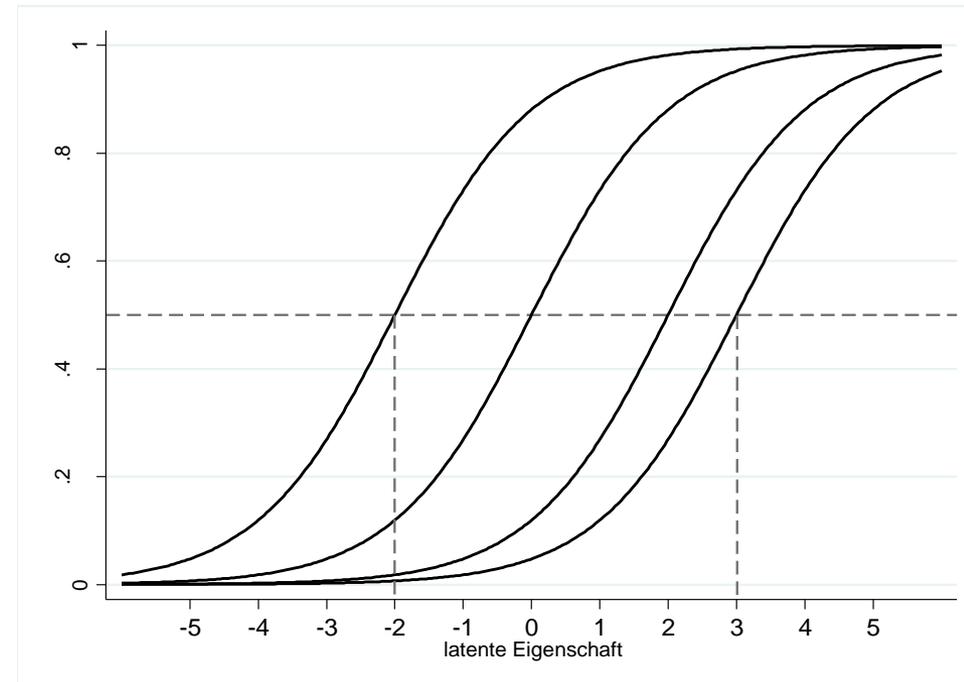
- Rasch-Modell für dichotome (1=richtig, 0=falsch) Testitems:

$$P(U_{ij} = 1 | \theta_i, \delta_j) = \frac{e^{\theta_i - \delta_j}}{1 + e^{\theta_i - \delta_j}}$$

mit  $\theta_i$  = Fähigkeit von Person  $i$

$\delta_j$  = Schwierigkeit von Item  $j$

- Die itemcharakteristischen Kurven verlaufen parallel (gleiche Trennschärfe).



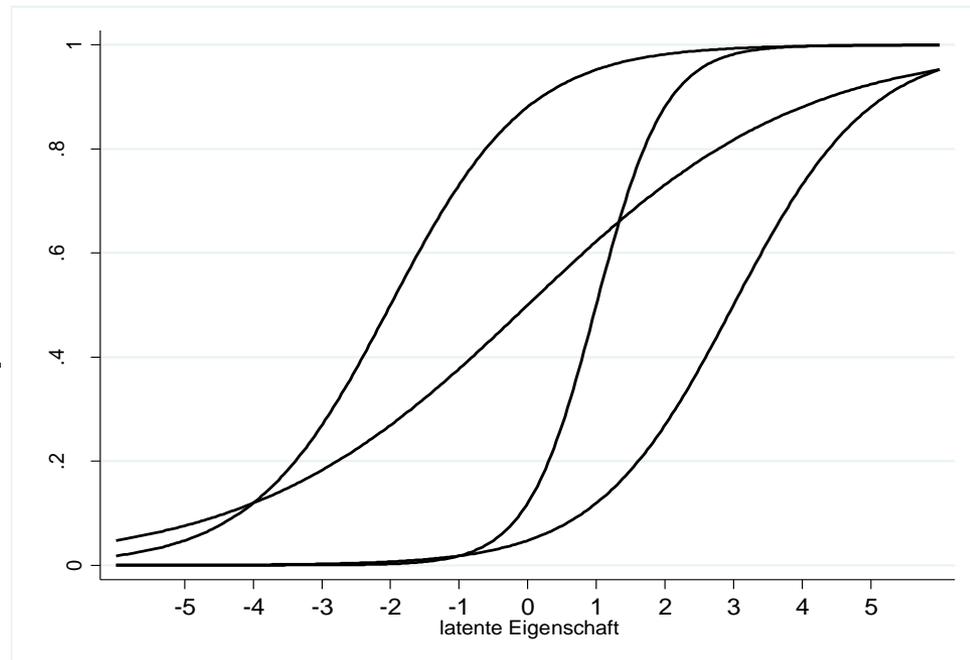
### Methoden: Item Response Theorie (IRT)

- Birnbaum-Modell für dichotome (1=richtig, 0=falsch) Testitems:

$$P(U_{ij} = 1 | \theta_i, \delta_j, \beta_j) = \frac{e^{\beta_j(\theta_i - \delta_j)}}{1 + e^{\beta_j(\theta_i - \delta_j)}}$$

mit  $\theta_i$  = Fähigkeit von Person  $i$   
 $\delta_j$  = Schwierigkeit von Item  $j$   
 $\beta_j$  = Trennschärfe von Item  $j$

- Die Rangfolge der Itemschwierigkeiten verändert sich über  $\theta$ .

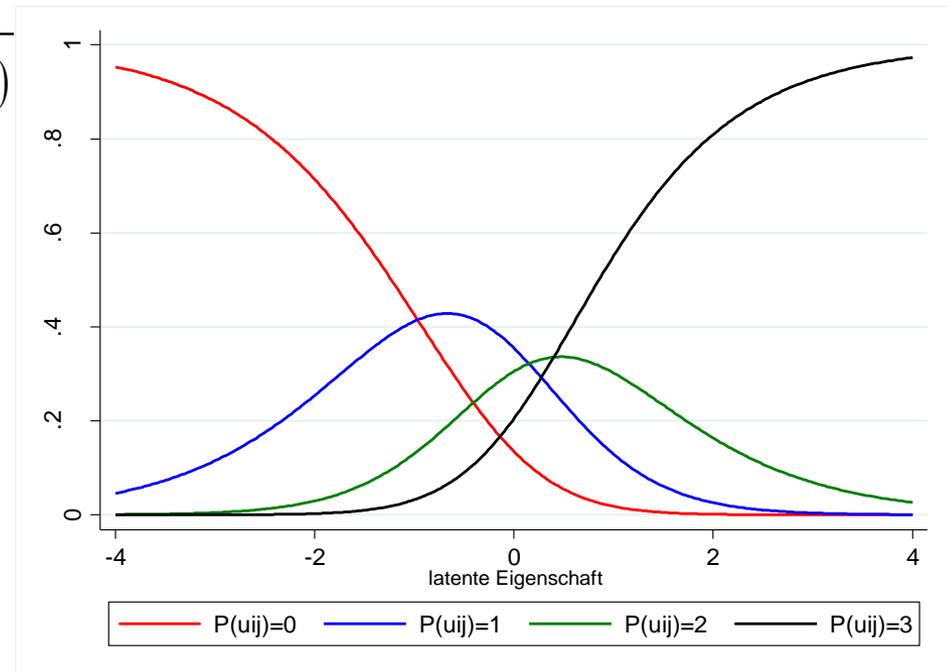


### Methoden: Item Response Theorie (IRT)

- **Partial Credit-Modell für ordinale Items mit  $m_j+1$  Antwortkategorien:**

$$P(u_{ij} = c | \theta_i, \delta_{jk}, \dots, \delta_{jm}) = \frac{e^{\sum_{k=0}^c (\theta_i - \delta_{jk})}}{\sum_{l=0}^{m_j} e^{\sum_{k=0}^l (\theta_i - \delta_{jk})}}$$

$$\sum_{k=0}^0 (\theta_i - \delta_{jk}) \equiv 0$$



Mit  $\theta_i$  = Fähigkeit von Person  $i$

$\delta_{jk}$  = Schwellenwert der  $k$ -ten Antwortstufe von Item  $j$

# Gliederung

1. Motivation
2. Begriffe, Konzepte und Methoden der Kompetenzforschung
- 3. Studiendesign**
  - Kompetenzmodell
  - Implementation und Ablauf
4. Ergebnisse
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. Fazit, offene Fragen und Ausblick

# Kompetenzstrukturmodell: Vorgehen

- **Ziel: Inhaltliche Definition und Dimensionierung von „Kompetenzen in Methoden der empirischen Sozialforschung“**
- **Vorgehen: teilsystematische Inhaltsanalyse von:**
  - Lehrbüchern
  - Curricula und Modulhandbüchern soziologischer Studiengänge
  - Lehrveranstaltungsplänen von Methodenkursen in der Soziologie
  - Empfehlungen der DGS-Methodensektion zur Methodenausbildung
  - Auszählung häufig genutzter Analyseverfahren in Fachzeitschriften (KZfSS, ZfS, BJS, MDA).

# Kompetenzstrukturmodell

- **Zwei Dimensionen:**
  - Datenerhebung
  - Datenanalyse / Statistik
  
- **Kognitive Prozesse („Bloom-Taxonomie“):**
  - jeweilige fachliche und methodische Inhalte (vgl. Krathwohl 2002)
    - kennen und verstehen
    - anwenden bzw. interpretieren
    - beurteilen/auswählen/gestalten

## Kompetenzstrukturmodell

Subdimension	Fachliche Inhalte	Kognitive Prozesse	leicht	mittel	schwer
Datenerhebung	Erhebungsdesigns	Kennen/Verstehen			
		Beurteilen/Auswählen			
	Stichprobenverfahren	Kennen/Verstehen			
Anwenden					
Beurteilen/Auswählen					
	Erhebungsmethoden	Kennen/Verstehen			
		Anwenden			
		Beurteilen/Auswählen/Gestalten			
	Mess- und Skalierungsverfahren	Kennen/Verstehen			
		Anwenden/Interpretieren			
		Beurteilen/Auswählen/Gestalten			
Datenanalyse / Statistik	Statistische Gesetze und Eigenschaften, insbesondere statistische Signifikanz	Kennen/Verstehen			
		Anwenden			
	Auswertungsverfahren und Technisches (Software)	Kennen/Verstehen			
		Anwenden			
		Beurteilen/Auswählen/Gestalten			
	Ergebnisse empirischer Analysen (uni-, bi-, multivariat)	Kennen/Verstehen			
		Anwenden/Interpretieren			
		Beurteilen/Auswählen/Gestalten			

# Erhebungsdesign

- **Selbstadministrierte PAPI-Befragung von Studierenden in Lehrveranstaltungen**
- **7 Universitäten (D/CH)**
- **Testheftdesign mit randomisierter Reihenfolge der Frageblöcke**
- **48 Items (32 pro Testheft) + subjektive Kompetenz, Hintergrundmerkmale**
- **N = 776 (davon 28 Dubletten und 11 weitere ausgeschlossene Fälle)**

# Erhebungsdesign

- **Selbstadministrierte PAPI-Befragung von Studierenden in Lehrveranstaltungen**
- **7 Universitäten (D/CH)**
- **Testheftdesign mit 10 verschiedenen Frageblöcken**
- **Erhebung von subjektive Kompetenz, Hintergrundmerkmale**
- **Erhebung von 28 Dubletten und 11 weitere ausgeschlossene Fälle)**

**Vielen Dank an alle, die uns bei der Datenerhebung unterstützt haben!**

# Fragetypen

Aufgaben mit freiem Antwortformat	Aufgaben mit gebundenem Antwortformat
10 Ergänzungsaufgaben	5 Zuordnungs- und Umordnungsaufgaben
12 Kurzaufsatzaufgaben	17 Single-Choice-Aufgaben
	4 Multiple-Choice-Aufgaben
$\Sigma = 22$	$\Sigma = 26$

# Gliederung

1. Motivation
2. Begriffe, Konzepte und Methoden der Kompetenzforschung
3. Studiendesign
  - Kompetenzmodell
  - Implementation und Ablauf
4. Ergebnisse
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. Fazit, offene Fragen und Ausblick

### Skalierung: Vorgehensweise

1. **Bewertung der Testitems (nicht/falsch gelöst, teilweise gelöst, vollständig gelöst)**
2. **Ausschluss von Dubletten und 11 weiteren Fällen**
3. **Schätzung von (eindimensionalen) IRT-Modellen für ordinale Items für jede der zwei inhaltlichen Dimensionen:**
  - Partial-Credit-Modell (PC, analog zu Rasch-Modell; Masters 1982)
  - Generalized Partial-Credit-Model (G-PC, analog zu Birnbaum-Modell; Muraki 1992)

# Skalierung: Vorgehensweise

### 4. Modelldiagnose:

- Korrelation zwischen Item und latenter Variable (Kompetenz)
- Outfit / Infit der Items
- Visuelle Inspektion der erwarteten und empirischen ICC
- Globaler Modelltest: 1PL- vs. 2PL-Modell (LR-Tests)
- Analysen zu Differential Item Functioning (DIF):
  - Geschlecht
  - Hauptfach/Nebenfach
  - Universität
- Dimensionalitätstests (Mehrdimensionale IRT-Modelle)

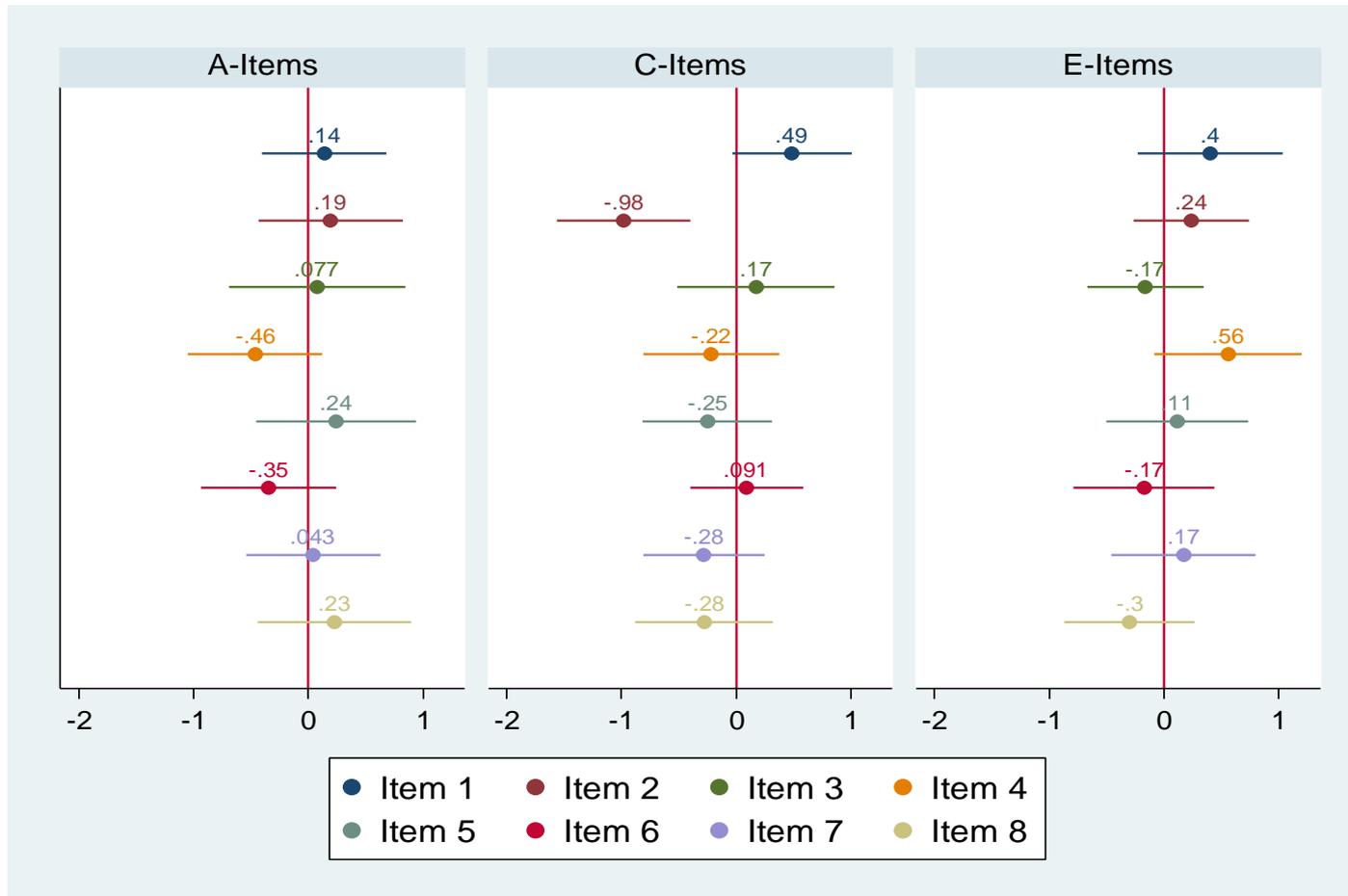
## Skalierung: Ergebnisse

	Datenerhebung	Datenanalyse
Erhobene Items	24	24
Ausgewählte Items	24	23
LR-Test PC vs. G-PC: $\chi^2$ (df)	92 (23) ***	297 (22) ***
BIC (PC)	16303	11572
BIC (G-PC)	16363	11420
Korrelation Item- $\theta$	alle > 0,35	alle > 0,34
Outfit/Infit	OK	OK
Gewähltes Modell	PC	G-PC
WLE Reliabilität	0,8	0,7
N	735	735

### Exkurs: Differential Item Functioning (DIF)

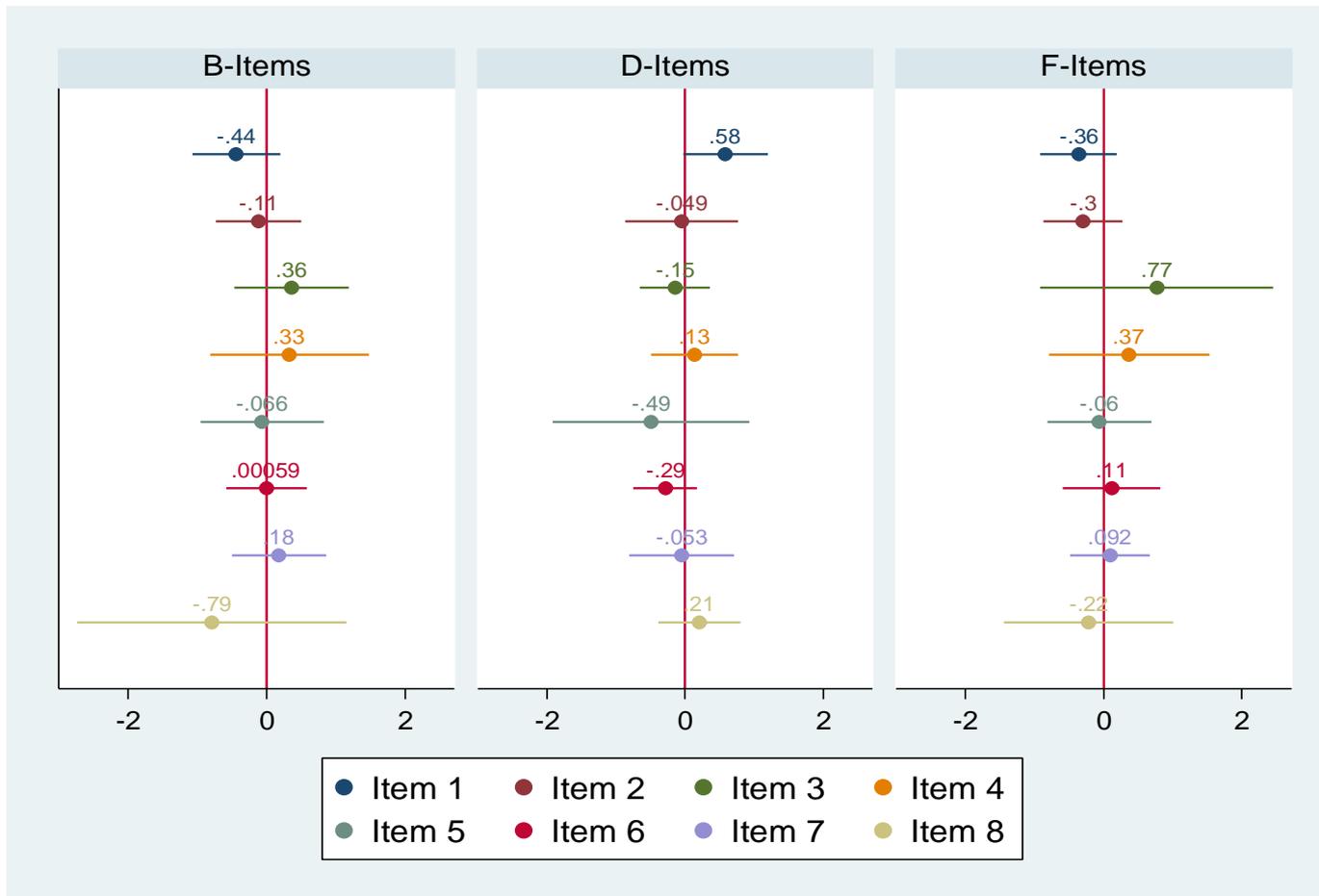
- **DIF tritt auf, wenn außer der latenten Variablen (der Kompetenz) auch andere Befragtenmerkmale einen Einfluss auf die Beantwortung eines Items haben.**
- **DIF kann auf „unfaire Testaufgaben“ hinweisen.**
- **Zumbo (1999) schlägt vor, logistische Regressionen der Testitems auf den Kompetenzwert und die interessierende DIF-Variable (sowie deren Interaktion, im Folgenden nicht dokumentiert) zu schätzen. Zeigt sich ein signifikanter und substantiell bedeutsamer partieller Effekt dieser Variablen, weist dies auf DIF hin.**

## DIF nach Geschlecht (Logitkoeffizienten und 99%-CI): Datenerhebung

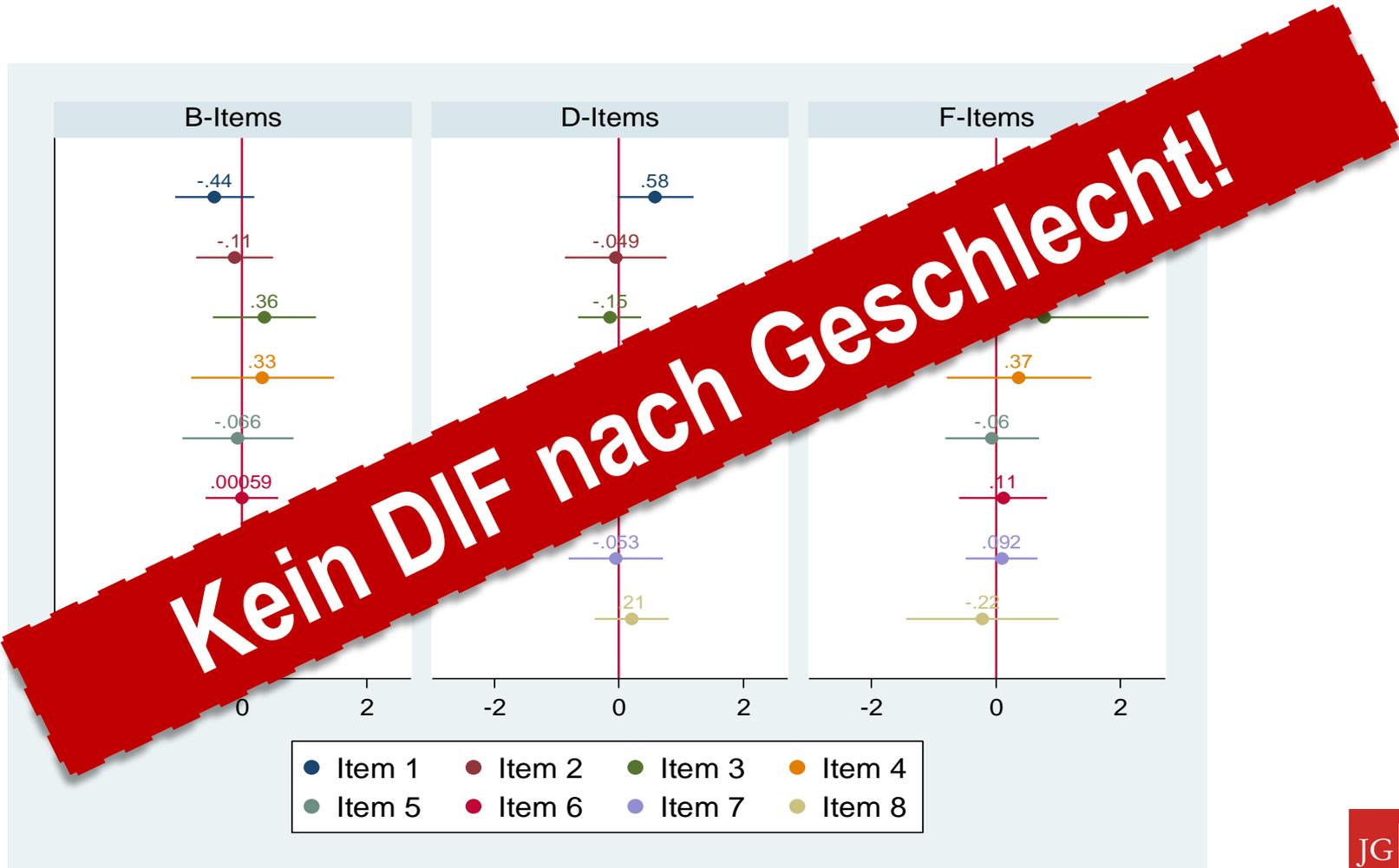


# 4. Ergebnisse: Skalierung der Kompetenzindikatoren

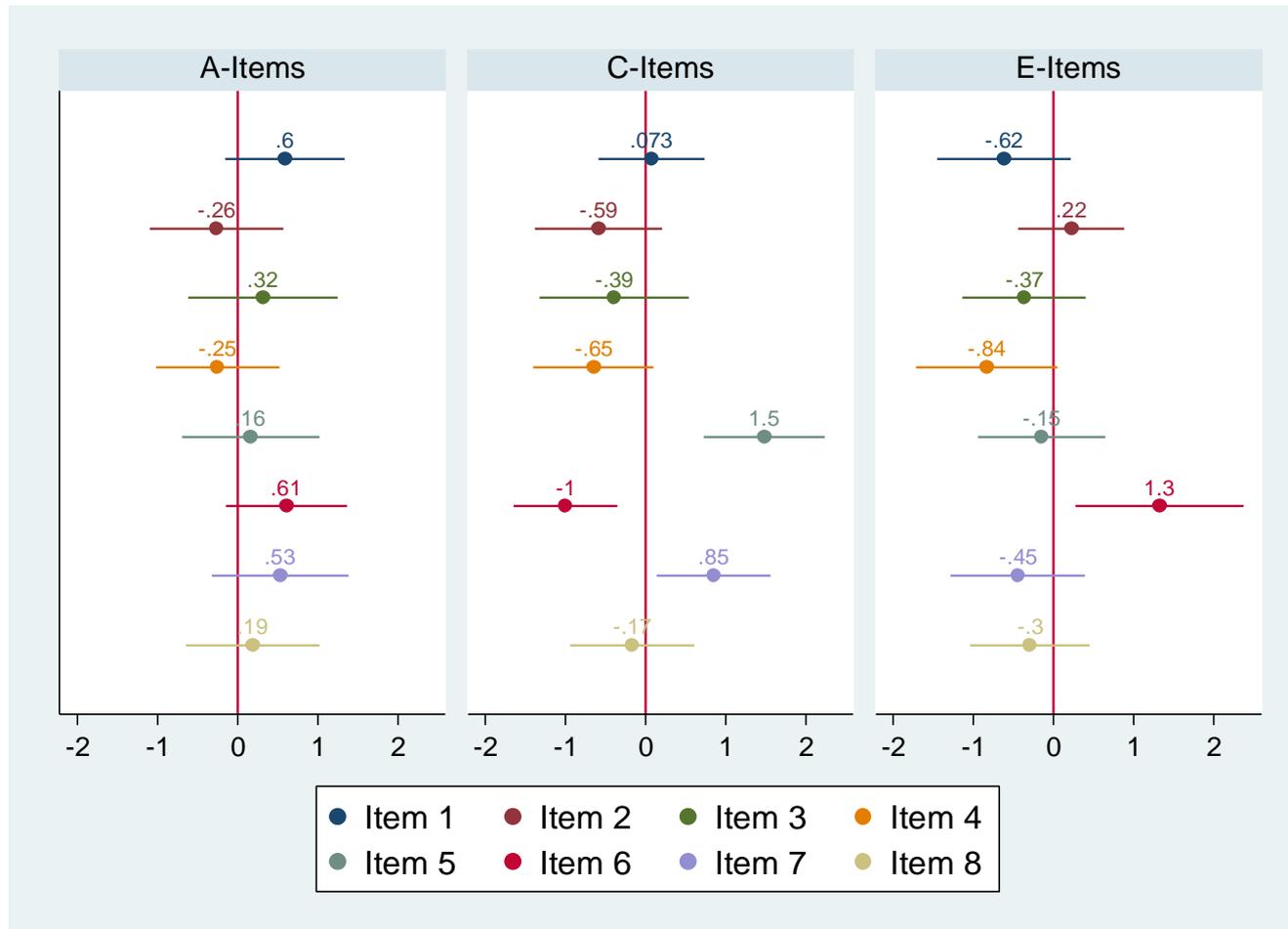
## DIF nach Geschlecht (Logitkoeffizienten und 99%-CI): Datenanalyse



## DIF nach Geschlecht (Logitkoeffizienten und 99%-CI): Datenanalyse



## DIF nach Uni (Logitkoeffizienten und 99%-CI): Beispiel



### DIF nach Hochschulstandort

- **DIF-Analysen für die einzelnen Uni-Standorte zeigen insgesamt, dass die Items an den jeweiligen Universitäten unterschiedlich funktionieren, wobei die Effekte allerdings substantiell nicht sehr stark sind.**
  - **Unsere These: Es ist unmöglich, Kompetenztests im Hochschulbereich ohne DIF zu entwickeln.**
- ➔ **Wie damit umgehen?**
- **Unsere Position: Pragmatisch bleiben! Sofern keine Hochschulen systematisch bei allen Items des Tests benachteiligt/bevorteilt werden, ist das i.O.**

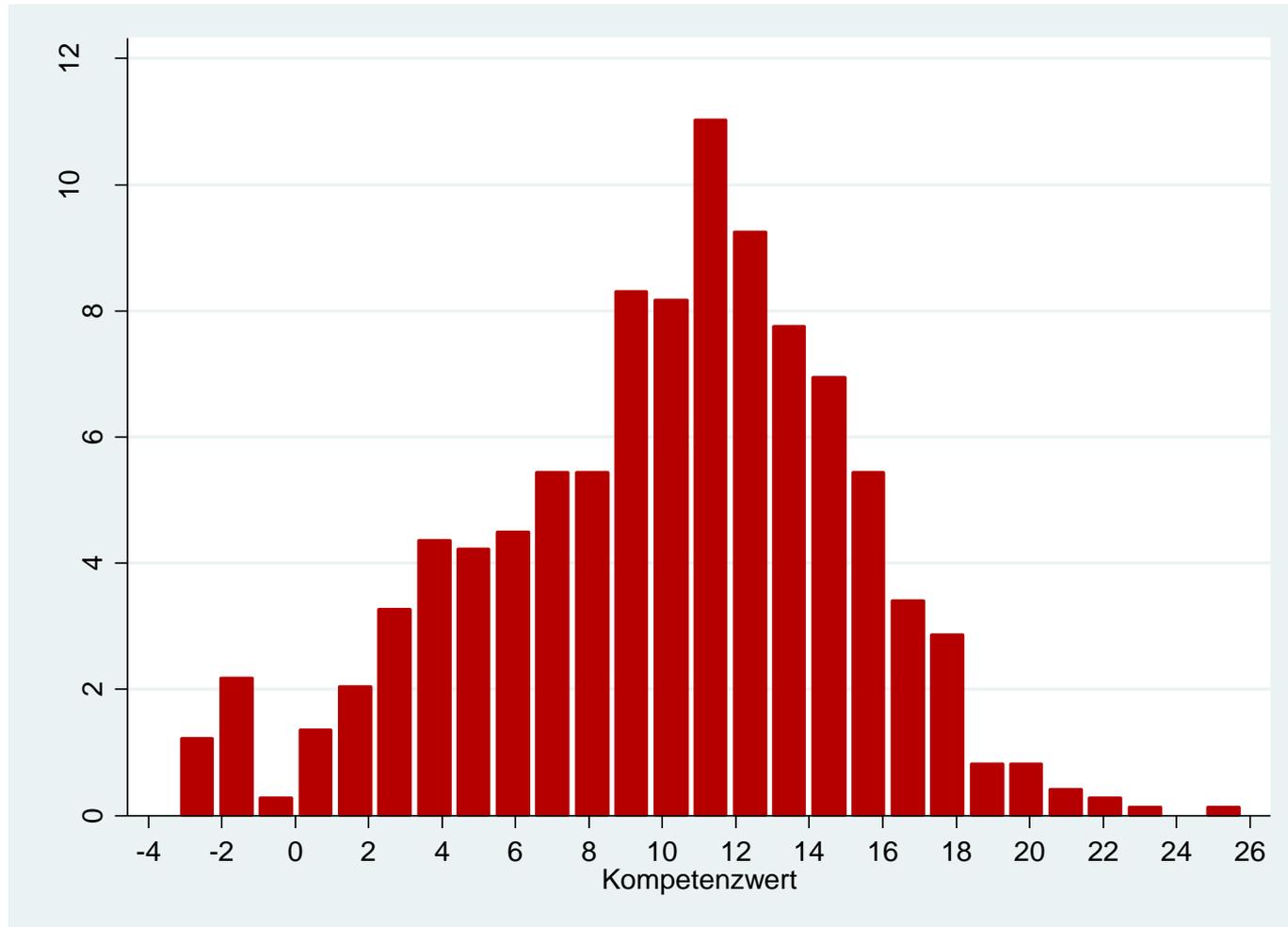




### Skalierung: Weitere Vorgehensweise

- **Für die folgenden inhaltliche Analysen: Zusammenfassen der zwei Subdimensionen zu einem Kompetenzindikator**
  - Eine Faktorenanalyse ergibt für die erste Hauptkomponente eine erklärte Varianz von 85 %.
  - Für die folgenden empirischen Analysen wurden die zwei Subdimensionen addiert und auf Mittelwert 10 und Standardabweichung 5 standardisiert.

## Globaler Kompetenzindikator



# Gliederung

1. Motivation
2. Begriffe, Konzepte und Methoden der Kompetenzforschung
3. Studiendesign
  - Kompetenzmodell
  - Implementation und Ablauf
4. Ergebnisse
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. Fazit, offene Fragen und Ausblick

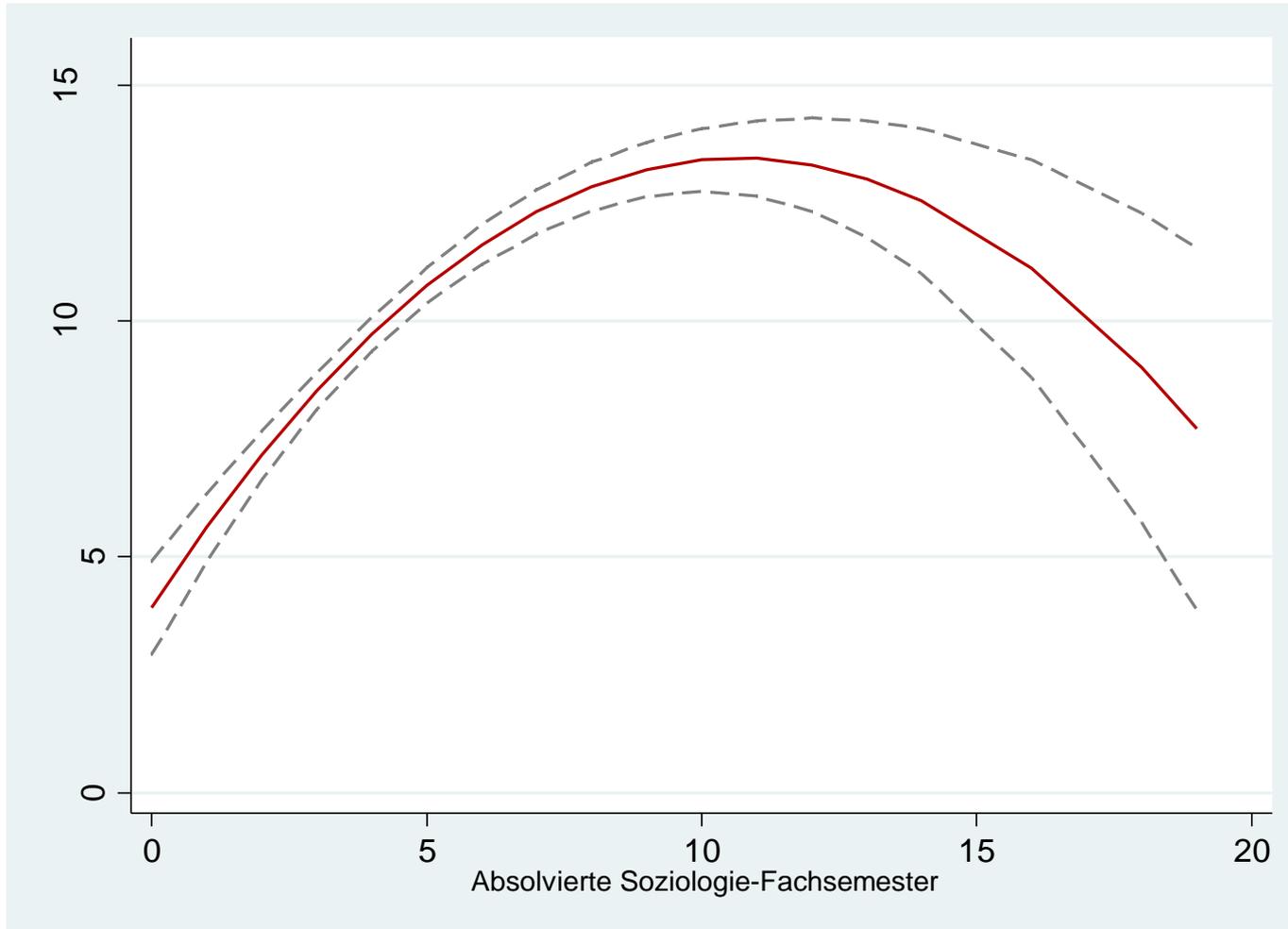
### Deskriptive Auswertungen

	Min	Max	Mittelwert	St.abw.	N
Kompetenzindikator	-3,21	25,83	10	5	735
Fachsemester in Sozialwissenschaften	0	19	4,92	2,84	728
Zahl der absolvierten Methodenkurse	0	4	2,08	1,26	679
Nebenfach (1 = ja)	0	1	0,33		729
Selbstwirksamkeit	1,33	7	4,31	1,05	701
Intrinsische fachliche Motivation	1	7	4,46	1,19	691
Abiturnote	1	4,6	2,41	0,56	718
Erwerbstätigkeit neben Studium (1 = ja)	0	1	0,53		706
Bildung der Eltern in Jahren	7	18	14,57	2,56	674
Migrationshintergrund (1 = ja)	0	1	0,17		688
Geschlecht (1 = weiblich)	0	1	0,62		703
Mühe beim Ausfüllen der Testitems (subjektiv)	1	7	4,13	1,55	703

### Determinanten der Methodenkompetenz (OLS/MI)

	Modell 1 (b)	Modell 2 (b)	Modell 3 (b)
Fachsemester (zentriert)	0,963 ***		
Fachsemester quadriert (zentriert)	-0,083 ***		
Zahl der absolvierten Methodenkurse			
Nebenfach (1 = ja)			
Selbstwirksamkeit			
Intrinsische (fachliche) Motivation			
Abiturnote			
Erwerbstätigkeit neben Studium (1 = ja)			
Bildung der Eltern in Jahren			
Migrationshintergrund (1 = ja)			
Geschlecht (1 = weiblich)			
Mühe beim Ausfüllen der Testitems			
Konstante	10,676 ***		
Korrigiertes R <sup>2</sup>	0,209		
N	735		

# Studienfortschritt (Fachsemester) und Kompetenz



### Determinanten der Methodenkompetenz (OLS/MI)

	Modell 1 (b)	Modell 2 (b)	Modell 3 (b)
Fachsemester (zentriert)	0,963 ***	0,367 ***	
Fachsemester quadriert (zentriert)	-0,083 ***	-0,047 ***	
Zahl der absolvierten Methodenkurse		2,036 ***	
Nebenfach (1 = ja)			
Selbstwirksamkeit			
Intrinsische (fachliche) Motivation			
Abiturnote			
Erwerbstätigkeit neben Studium (1 = ja)			
Bildung der Eltern in Jahren			
Migrationshintergrund (1 = ja)			
Geschlecht (1 = weiblich)			
Mühe beim Ausfüllen der Testitems			
Konstante	10,676 ***	6,202 ***	
Korrigiertes R <sup>2</sup>	0,209	0,378	
N	735	735	

### Determinanten der Methodenkompetenz (OLS/MI)

	Modell 1 (b)	Modell 2 (b)	Modell 3 (b)
Fachsemester (zentriert)	0,963 ***	0,367 ***	0,486 ***
Fachsemester quadriert (zentriert)	-0,083 ***	-0,047 ***	-0,051 ***
Zahl der absolvierten Methodenkurse		2,036 ***	1,253 ***
Nebenfach (1 = ja)			-2,516 ***
Selbstwirksamkeit			0,378 **
Intrinsische (fachliche) Motivation			0,086
Abiturnote			-1,465 ***
Erwerbstätigkeit neben Studium (1 = ja)			0,034
Bildung der Eltern in Jahren			-0,050
Migrationshintergrund (1 = ja)			-0,632 +
Geschlecht (1 = weiblich)			-0,982 **
Mühe beim Ausfüllen der Testitems			0,526 ***
Konstante	10,676 ***	6,202 ***	9,443 ***
Korrigiertes R <sup>2</sup>	0,209	0,378	0,482
N	735	735	735

### Ergebnisse der Regressionsanalysen

- **Bezüglich unseres Kompetenzindikators zeigt sich eine hohe Konstruktvalidität mit getrennt erhobenen Außenkriterien.**
- **Die Analysen replizieren nahezu exakt die Ergebnisse der ersten Pilotstudie („soziologische Kompetenzen“).**
- **Zum Geschlechtseffekt:**
  - substanzuell?
  - unterschiedliches Ankreuzverhalten?
  - DIF?

# Ergebnisse der Regressionsanalysen

- **Bezüglich unseres Kompetenzindikators zeigt sich eine hohe Konstruktvalidität mit getrennt erhobenen Außenkriterien.**
- **Die Analysen replizieren nahezu exakt die Ergebnisse der ersten Pilotstudie („soziologische Kompetenzen“).**
- **Zum Geschlechtseffekt:**
  - substanziiell!?
  - unterschiedliches Ankreuzverhalten?
    - Nur schwacher Zusammenhang zwischen Zahl der nicht ausgefüllten Items und Geschlecht
  - DIF?
    - Kein DIF nach Geschlecht

# Unterschiedliche Kompetenz nach Hochschule?

- Fallzahlen nach Hochschule:

**anonymisiert**

### Mehrebenenmodelle: Determinanten der Methodenkompetenz

	Modell 1 (b)	Modell 2 (b)
Fachsemester (zentriert)		
Fachsemester quadriert (zentriert)		
Ebene 1-Kontrollvariablen		
Konstante		
Varianz Fachsemester		
Varianz Konstante		
Varianz Residuum		
Intraclassen		
N		

**anonymisiert**

## Effekt des Studienfortschritts nach Hochschule

**anonymisiert**

# Gliederung

1. Motivation
2. Begriffe, Konzepte und Methoden der Kompetenzforschung
3. Studiendesign
  - Kompetenzmodell
  - Implementation und Ablauf
4. Ergebnisse
  - Skalierung der Kompetenzindikatoren
  - Ausgewählte inhaltliche Analysen: Konstruktvalidität und Analysepotenzial
5. **Fazit, offene Fragen und Ausblick**

### Fazit

- **Eine direkte Übertragung der kompetenzdiagnostischen Methoden aus dem schulischen Bereich auf den Hochschulbereich ist mit einigen Hürden verbunden. Aber:**
  - Die Motivation / der Bedarf für solche Unterfangen ist auf jeden Fall gegeben.
  - Methodisch „liegt auf dem Tisch“, was zu leisten ist.
- **Positive Erfahrungen:**
  - Erhebungspraktisch gut durchführbar.
  - Skalierung der Konstrukte funktioniert gut.
  - Aussagekräftige inhaltliche Analysen/gute Konstruktvalidität.

# Problembereiche (I)

- **Kompetenzstrukturmodelle:**
  - Im Gegensatz zum schulischen Bereich kein einheitlicher Lehrkanon an Hochschulen
  - Wie generische (z.B. analytisches Denken) von fachspezifischen Kompetenzen trennen (und messbar machen)?
  - Professionspolitisch konfliktträchtig
- **Messung von „langwierigen Denkprozessen“ mit standardisierten Instrumenten?**
- **Hochschulevaluation:**
  - Wie kausale Effekte von Hochschulen (oder Curricula usw.) isolieren?
  - Derartige Probleme sind aber kein Spezifikum von Kompetenzmessungen.

### Problembereiche (II)

- **Auswahl der Untersuchungseinheiten und Zeitpunkte der Messungen im Studienverlauf?**
  - Ideal: Prior Ability messen und dann Panelstudie bis Verlassen der Hochschule?!
- **DIF nach (Neben)Fächern und Universität?**
- **Mehr Items, mehr Fälle und grundsätzlich mehr Ressourcen nötig.**
  - Schulleistungsstudien: Mehrere Pilotierungsrunden mit mehreren zehntausend Probanden.

**Vielen Dank!**

[felix.wolter@uni-mainz.de](mailto:felix.wolter@uni-mainz.de)

[juergen.schiener@uni-mainz.de](mailto:juergen.schiener@uni-mainz.de)

## Literatur

- Birnbaum, Allan 1968: Some Latent Trait Models, in: Lord, Frederic M./Novick, Melvin R. (Hrsg.): **Statistical Theories of Mental Test Scores**, Reading, MA: Addison-Wesley: S. 395–479.
- Hartig, Johannes/Klieme, Eckhard 2006: **Kompetenz und Kompetenzdiagnostik**, in: Schweizer, Karl (Hrsg.): **Leistung und Leistungsdiagnostik**, Heidelberg: Springer: S. 127–143.
- Kiefer, Thomas/Robitzsch, Alexander/Wu, Margaret 2014: **Package tam**, WWW-Dokument: <http://cran.r-project.org/web/packages/TAM/TAM.pdf>, 18.01.2014.
- Klieme, Eckhard/Leutner, Detlev 2006: **Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG**, in: **Zeitschrift für Pädagogik** 52 (6): S. 876–903.
- Krathwohl, David R. 2002: **A Revision of Bloom's Taxonomy: An Overview**, in: **Theory into Practice** 41 (4): S. 212–218.
- Masters, G. N. 1982: **A Rasch Model for Partial Credit Scoring**, **Psychometrika** 47, 149–174.
- Muraki, E. 1992: **A Generalized Partial Credit Model: Application of an EM Algorithm**, **Applied Psychological Measurement** 19, 159–176.
- Pohl, Steffi/Carstensen, Claus H. 2012: **NEPS Technical Report – Scaling the Data of the Competence Tests (NEPS Working Paper No. 14)**, Bamberg: University of Bamberg.
- Rasch, Georg 1960: **Probabilistic Models for some Intelligence and Attainment Tests**, Kopenhagen: The Danish Institute for Educational Research.
- Roth, Heinrich 1971: **Pädagogische Anthropologie. Band 2: Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik**, Hannover: Schroedel.
- Wolter, Felix/Schiener, Jürgen 2014: **Auf dem Weg zum „Hochschul-PISA“? Zur Messung „soziologischer Kompetenzen“**, **Soziale Welt** 65 (1).
- Zumbo, B. D. 1999: **A Handbook on the Theory and Methods of Differential Item Functioning (DIF)**, Ottawa: Department of National Defense.

## Beispielfragen

**FC4** Berechnen Sie Modus, Median und arithmetisches Mittel für diese Gruppe von Studenten:

<b>Student:</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Alter:</b>	<b>20</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>

a) Modus:

b) Median:

c) Arithmetisches Mittel:

**FC4** Sie möchten eine quantitative Erhebung unter Heroinabhängigen in Ihrem Heimatort durchführen und dabei möglichst alle Personen der Zielgruppe erfassen. Welches Stichprobenverfahren bietet sich an?

- Einwohnermeldeamtstichprobe  1
- Schneeballverfahren  2
- Zufallsbasierte Klumpenstichprobe  3
- Reine Zufallsauswahl  4

**FD2** Vervollständigen Sie die absoluten Häufigkeiten in der folgenden Kreuztabelle, so dass ein auf Chi-Quadrat basiertes Assoziationsmaß einen Wert von 1 (perfekter Zusammenhang) einnimmt.

	kein Abitur	Abitur	Randsumme
kein Zeitungsabonnent			
Zeitungsabonnent			
Randsumme			100

## Objektive Kompetenz, subjektive Kompetenz und Noten

- **Subjektive Kompetenz:**

**F13** Wenn Sie Ihre persönlichen Kompetenzen in den Methoden der empirischen Sozialforschung einmal subjektiv bewerten würden: Wie gut kennen Sie sich in den unten aufgeführten Bereichen aus?

a) **Methoden der empirischen Sozialforschung insgesamt:**

kenne mich überhaupt nicht aus ←————→ kenne mich sehr gut aus

1    2    3    4    5    6    7

b) **Methoden der Datenerhebung:**

1    2    3    4    5    6    7

c) **Methoden der Datenanalyse / Statistik:**

1    2    3    4    5    6    7

- **Noten: Durchschnittsnote aus bis zu vier Methodenveranstaltungen**

## Objektive Kompetenz, subjektive Kompetenz und Noten

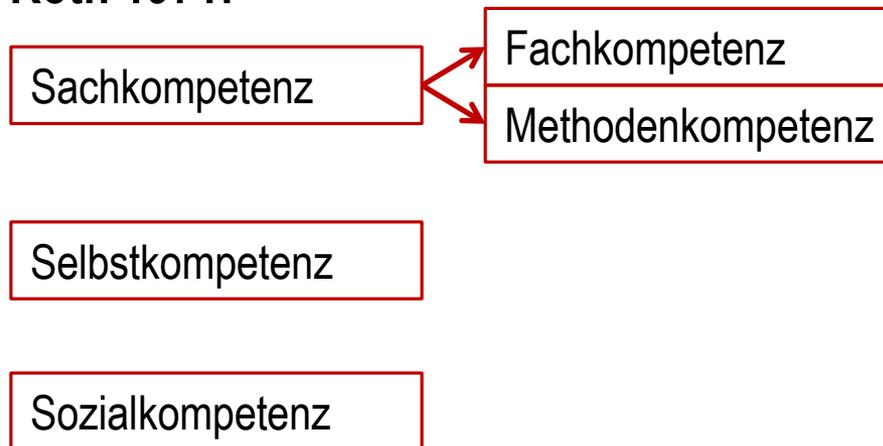
- **Korrelationsmatrix:**

	Objektive Kompetenz	Subjektive Kompetenz
Subjektive Kompetenz	0,531*** N=698	
Noten	-0,362** N=526	-0,299*** N=512

# Zum Kompetenzbegriff

- **Gängige Definition von „Kompetenzen“:**
  - „*kontextspezifische kognitive Leistungsdispositionen*, die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme/Leutner 2006: 879, Herv. i. Orig.).
  - Handeln können („Can-Do-Aussagen“)
  - Erlernbarkeit (im Unterschied zu Konstrukten wie „Intelligenz“)

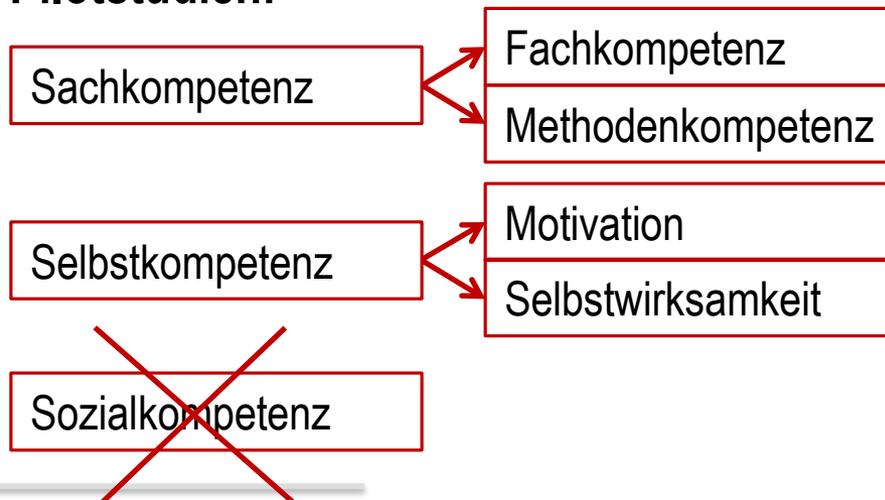
- **Roth 1971:**



# Zum Kompetenzbegriff

- **Gängige Definition von „Kompetenzen“:**
  - „*kontextspezifische kognitive Leistungsdispositionen*, die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme/Leutner 2006: 879, Herv. i. Orig.).
  - Handeln können („Can-Do-Aussagen“)
  - Erlernbarkeit (im Unterschied zu Konstrukten wie „Intelligenz“)

- **Pilotstudien:**



### Testheftdesign von Studie 2

Testheft		1	2	3	4	5	6
Frageblöcke	Position 1	AB	CD	AB	CD	EF	EF
	Position 2	CD	EF	EF	AB	CD	AB

- Subdimension Datenerhebung: Je 8 Items in Block A, C und E.
- Subdimension Datenanalyse/Statistik: Je 8 Items in Block B, D und F.