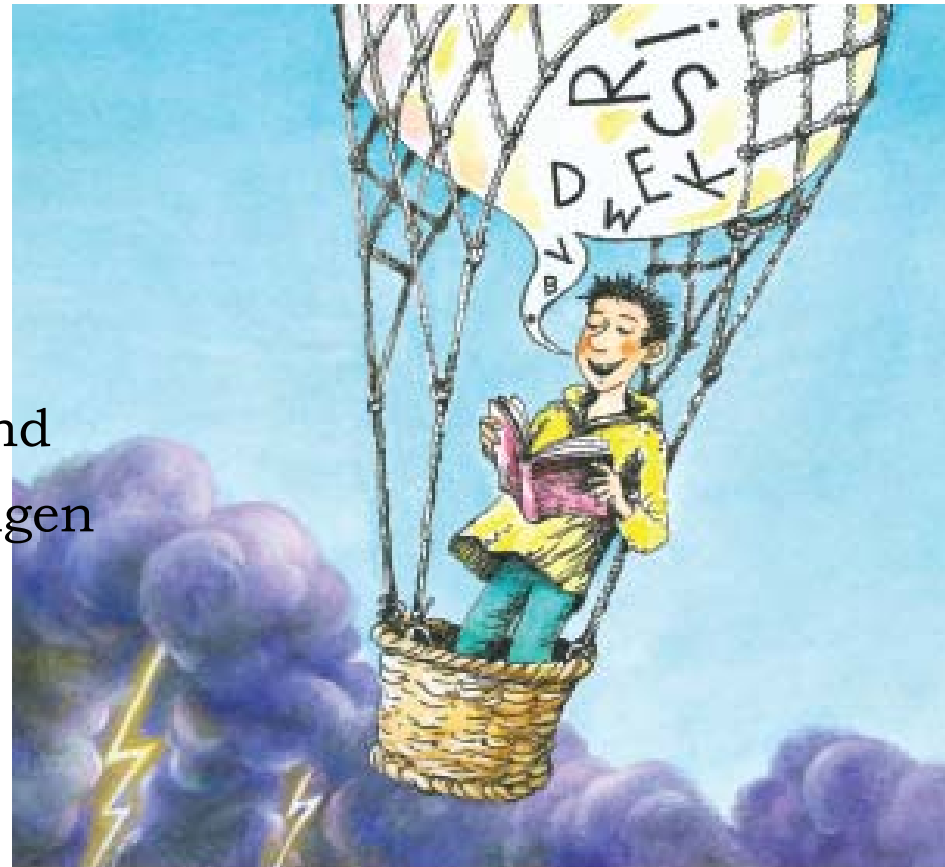




Universität Hamburg

Heiner Willenberg

Validität,  
Reliabilität und  
Objektivierungen



in der deutsch-  
didaktischen  
Empirie

# Was ist DESI?

- 1. Deutsch – Englisch – Schülerleistungen – International
- 2. Im Auftrag der KMK: 2002 – 2005  
Untersuchung von ca. 11.000 Schüler/innen am Anfang und Ende der 9. Klasse (alle Bildungsgänge, alle Bundesländer)
- 3. In beiden Fächern die volle Lehrplan-  
“breite“

# DESI Kompetenzmodelle

<b>LESEN</b>	<b>WORTSCHATZ</b>	<b>ARGUMENTATION</b>
Vier Kompetenzstufen (Sechs Lesetätigkeiten)	Drei Kompetenzstufen	Drei Kompetenzstufen
<b>SCHREIBEN</b>	<b>SPRACHBEWUSSTHEIT</b>	<b>RECHTSCHREIBEN</b>
Vier Kompetenzstufen	Drei Kompetenzstufen	Drei Kompetenzstufen

# DESI-Berichte

- 1. Heiner Willenberg (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Baltmannsweiler: Schneider, 239 Seiten (2007)
- 2. Bärbel Beck & Eckhard Klieme (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messungen. DESI-Studie. Weinheim: Beltz, 319 Seiten (2007)

# Validität (interne)

- 1. Interne Validität:
  - „Das Ausmaß, in dem es gelingt, [störende] Drittvariablen zu kontrollieren
  - und in dem ein Kausalschluss von der unabhängigen auf die abhängige [untersuchte] Variable gerechtfertigt ist.“  
(Schreier, 187)
  - > Designpläne
  - > Kovarianzen
- DESI
  - > Designpläne
  - > Kovarianzen

# Validität (intern): Designpläne (Schreier)

	Vorunter- suchung	Treatment	Zwischen- Phase(n)	Nachunter- suchung	Kommentar
Einmal	#	ja	#	ja	Zu vermeiden
Prä-Post	ja	Ja	#	ja	klassisch
<i>Verum- /</i>	<i>Ja</i>	<i>Ja</i>		<i>ja</i>	<i>besser</i>
<i>Kontroll- gruppe</i>	<i>Ja</i>	<i>Nein</i>		<i>ja</i>	<i>besser</i>
+ Zwischen- phasen	ja	ja	ja / ja	ja	Noch besser

# Validität (intern)

## DESI-Design

	<b>Vorunter- suchung</b>	Treatment / <b>Zwischenphase</b>	Nachunter- suchung
Prä / Post	Erster Mess- zeitpunkt	Normaler Unterricht in einem Jahr	Zweiter Mess- zeitpunkt
<i>Kontrolle:</i>	<i>Repäsentative Stichprobe</i>		
<i>Rotation der Aufgaben</i>	<i>an beiden</i>		<i>Messzeit- punkten</i>
<i>Zunächst nur Bestandsauf- nahme</i>	<i>Dann Hypothesen:</i>	<i>Zusammenhang von Leistungen und Kovarianzen</i>	<i>im Fragebogen</i>

# Validität (intern)

## Kovarianzuntersuchungen

### *Beispiel: Lesen*

- Intelligenz
- *Leseinteresse*
- Kulturelle Familieninteressen
- Unterrichtsstil
- Bildungsgänge
- Geschlecht / Sprachherkunft
- KFT (Kogn. Fähigk. Test)
- Lesegeschwindigkeit
- *Leseinteresse (Skala mit 13 Items)*
- Fragebogen (Hisei) Eltern
- Fragebögen für Schüler,
- Wichtigkeit sprachlicher Aspekte (10 Items) / Leseinitiativen
- Stichprobe
- Jeweils erfasst:
  - Deutsch, zweisprachig, andere Erstsprache als Deutsch

# Zwei Skalen

- Leseinteresse (13 Items)
  - 1. Weil mir das Lesen Spaß macht, würde ich nur ungern darauf verzichten.
  - 2. Ich lese in meiner Freizeit.
  - 3. Wenn ich lese, vergesse ich manchmal alles um mich herum.
- Sprachliche Fähigkeiten (10 Items)
  - *Wie wichtig ist deinem Deutschlehrer, deiner Deutschlehrerin Folgendes: (für euch)*
  - Rechtschreibung und Zeichensetzung beherrschen
  - Die wichtigsten Inhalte eines Textes zusammenfassen
  - Überzeugend argumentieren
  - Wichtige gramm. Begriffe kennen
  - Im Gespräch dem Partner zuhören und auf ihn eingehen

## Kovarianz 1: Korrelationen (DESI)

alle im Bereich 0,01 signifikant

	KFT	Lese- gesch w.	Lese- Inter- esse	Hisei	Mut- ter- spra- che	Spr. Fertig- - keiten	Ge- schlecht	Lernsit: Lese- verste- hen
Lese- Komp . LK	<b>,563</b>	,336	<b>,318</b>	<b>,294</b>	<b>-.230</b>	<b>,168</b>	<b>,095</b> (w)	-,021
Lese- inter.	<u>,255</u>	,256		<u>,167</u>	<b>-,073</b>	<u>,197</u>	<b>,293</b> (w)	<b>,093</b>

# Deutung der Korrelationen 1

- **1. Unabänderliche Einflüsse**
- Intelligenz (.563), Hisei (.294), Muttersprache (-.230), Geschlecht (.095)
- [Highest International Socio-Economic Index, HISEI]
- **2. Veränderbare Einflüsse**
- Erzeugung von Leseinteresse (.318), Wert auf sprachliche Fähigkeiten (.168)
- Unklar: Lesegeschwindigkeit (trainierbar? .336)?
- Initiativen zur Leseförderung: Warum negativ (-.021)

## Deutung der Korrelationen 2

- Wenn Leseinteresse vorhanden ist,
- Erhöht sich der Einfluss: des (weiblichen) Geschlechts (.095 auf .293), **das Gewicht des Anspruchs auf sprachliche Fähigkeiten (.168 auf .197)**
- Verringert sich der Einfluss: der Intelligenz (.563 auf .255), der sozialen Güter (.294 auf .167), der Lesegeschwindigkeit (.336 auf .256), der negative Einfluss der fremden Muttersprache (-.230 auf -.073)
- Werden die Leseinitiativen der Lehrerschaft milde positiv (-.021 auf .093)
- D.h. das Leseinteresse gleicht vieles aus, es ist aber weiblich bestimmt!

## Kovarianz 2: Multiple Regression

- Die Regressionsanalyse versucht, die Zusammenhänge zu optimieren und eventuell Korrelationen auszuschließen.
- Hier ist das Gewicht des Geschlechts gering, d.h. da müssen noch andere Kovarianzen vorhanden sein – u.a. in den vier Bildungsgängen.
- Man sieht bei ihr u.a., wie hoch die Varianzaufklärung der durch die gewählten Items ist (Streuung):
- Korrigiertes R Quadrat: hier .35 [gutes Maß]
- Prüft verdeckte Zusammenhänge, ob Items dasselbe messen und also nicht getrennt dargestellt werden sollten (Kollinearität). Ist hier nicht der Fall.

## Kovarianz 2: Regressionsanalyse: Erklärungen für die Lesekompetenz (2. TZP) (r Quadrat:.354)

	<b>Beta (Einfluss)</b>	Signifikanz.	Kollinearität (> .1)
<b>KFT: Intelligenztest (1)</b>	<b>,404</b>	,000	,813
<b>Leseinteresse T2 (SFB) (3)</b>	<b>,163</b>	,000	,853
<b>Höchster ISEI in Familie (Elternfragebogen) (4)</b>	<b>,107</b>	,000	,891
<u>Lesegeschwindigkeit (2)</u>	<b>,101</b>	,000	,860
<b>Wichtigkeit sprachlicher Fertigkeiten (SFB) (6)</b>	<b>,076</b>	,000	,818
<b>Muttersprache (5)</b>	<b>-,067</b>	,000	,954
<b>Lernsituationen im DU: Leseverstehen (SFB) (8)</b>	<b>-,058</b>	,000	,827
<b>Geschlecht (7)</b>	<b>-.009</b>	,000	,903

# Externe Validität (2)

- „Verallgemeinerbarkeit der Ergebnisse einer Untersuchung auf
  - andere Personen,
  - Situationen sowie
  - andere Formen der Erfassung der interessierenden Variablen.“
  - (Schreier)
- DESI
  - Nur 9. Klassen repräsentativ
  - Nur Fähigkeiten im Unterricht
  - Ergebnisse in den meisten Teilgebieten des Unterrichts
  - (keine Beobachtung freier Gespräche, echter Mündlichkeit)

# Konstruktvalidität (3)

## (nach Bortz & Döring)

- „Ein Test ist konstruktvalide, wenn aus dem zu messenden Zielkonstrukt Hypothesen ableitbar sind, die anhand der Testwerte bestätigt werden können.“ (186)
- Für die theoretische Grundlage müssen alle verfügbaren bisherigen Theorien geprüft bzw. herangezogen werden. Und es muss gefragt werden, ob sie Abstufungen ermöglichen?
- DESI war zunächst nur eine Bestandsaufnahme.
- Kausalzusammenhänge wurden hergestellt zu den Aspekten in der Kovarianz: Intelligenz, Elternhaus, Leseinteresse, Geschlecht, Sprache, Unterricht
- Für die Lesetheorie z.B. wurden alle vorliegenden Leseforschungen geprüft. Die Abstufungen nach Kintsch; Kirsch; Mentales Modell und Willenberg wurden aufgenommen.

# Abnehmervvalidität (4)

## (nach Wahl / Bortz & Döring)

- Die potentiellen Abnehmer müssen mit den Ergebnissen bessere Einsichten und Anwendbarkeiten erhalten, als sie es ohne die Untersuchung gehabt hätten. (Wahl)
  - „Wenn die Entscheidungen und Vorhersagen [bzw. die Beobachtungen], die auf der Basis des Tests getroffen werden tauglicher sind als die Entscheidungen und Vorhersagen, die ohne den Test möglich wären.“  
Bortz & Döring, 187
  - Die Beschreibungen der Lesestufen erlauben es, Unterrichtsprozesse genauer zu beobachten und v.a. auch Verbindungen zur Lese- und Verstehensgüte herzustellen.
-

# Verbindung qualitativer und quantitativer Methoden

- Modell 1:
- QUAL. kontinuierliche Sammlung
- QUANT. beider Datensorten (*z.B. Unterrichtsbeobachtungen illustrierend / Tests zur Lesefähigkeit*)
- Modell 2:
- QUANT. Welle 1 Welle 2 Welle 3
- (*Längsschnitt-Testung z.B. in Klassen: Leseleistung*)
- QUAL. \_\_\_\_\_
- Kontinuierliche Feldforschung (*z.B. Unterrichtsbeobachtungen: Schlüsselstellen für Leseleistungen*)

## Nach Miles & Huberman, 1994, 41 s.a. Flick: Triangulation

- Modell 3:
- QUAL: Exploration > QUANT. Fragebogen > QUAL.
- Vertiefung und Überprüfung von Ergebnissen
- *Unterrichtsbeobachtung versch. Klassen / Leseinteresse erfragt / Zusammenhänge interpretiert anhand von Transkripten*
- Modell 4:
- QUANT. Umfrage > QUAL. Feldstudie > QUANT. Experiment
- *Fragebogen zum Leseinteresse > Beobachtungen > zufallsausgesuchte Schülerprobanden werden getestet*

# Ihre Pläne?

- Wie ordnen sie sich ein?
- 1. Interne Validität und Ausschluss der Störvariablen?
- 2. Konstruktvalidität: Hypothesen / Theorie mit Abstufungen?
- 3. Kombination qualitativer und quantitativer Methoden?