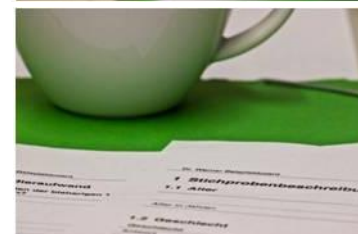




Messmodelle in der Lehrevaluation

Einfluss der hierarchischen Struktur auf
den Modellfit für Instrumente der
Lehrevaluation

1. Lehrevaluation an der FSU Jena
2. Messmodell für PELVE
 1. Modellfit bei Verletzung der iid-Annahme
 2. Aggregierte Datenebene vs. itembasierte Auswertung
3. Zusammenfassung & Ausblick



1. Lehrevaluation an der FSU Jena

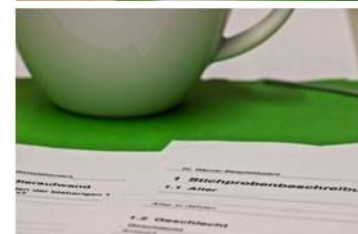
2. Messmodell für PELVE

1. Modellfit bei Verletzung der iid-Annahme

2. Aggregierte Datenebene vs. itembasierte

Auswertung

3. Zusammenfassung & Ausblick



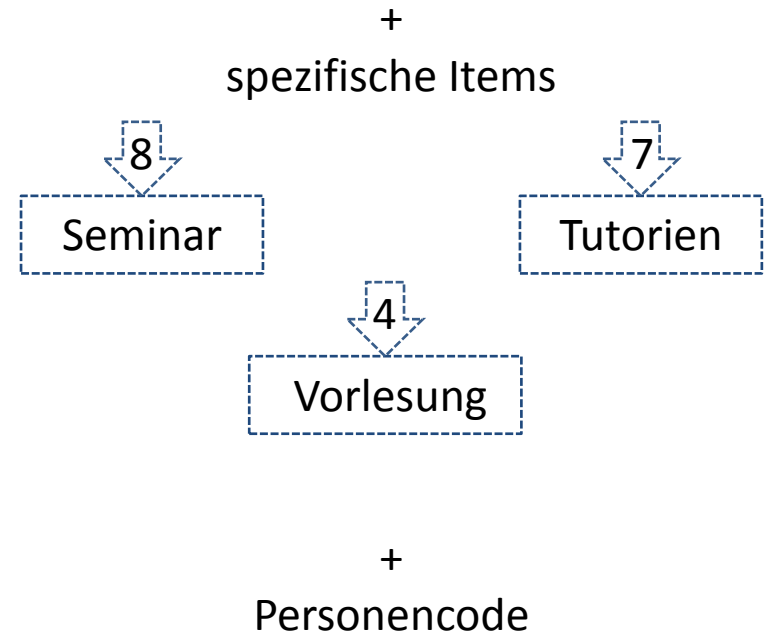
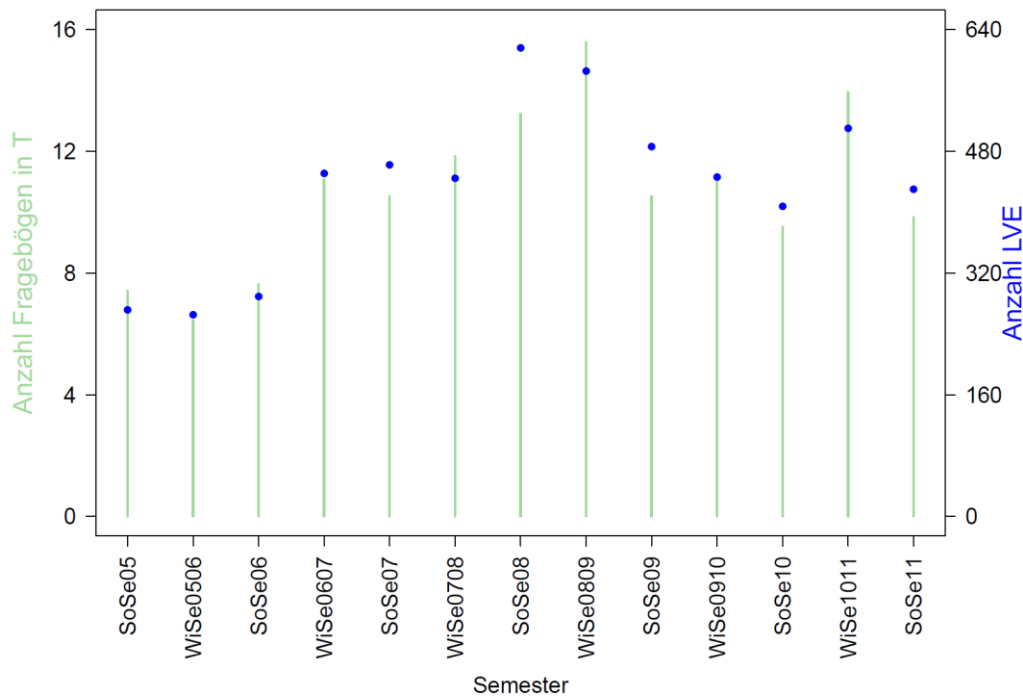
1. Lehrevaluation an der FSU



PELVE (Prozess und Ergebnisorientierte Lehrveranstaltungsevaluation)

35 Items als paper-pencil & online Version

LVE an der FSU



1. Evaluationsinstrument PELVE

Prozess- und Ergebnisorientierte Lehrveranstaltungsevaluation

Qualität einer Lehrveranstaltung
(Zusammenwirken von Prozess- und Ergebnisvariablen)

Prozessvariablen

(Prinzipiell veränderbar, bilden den Prozess der Lehrveranstaltung ab)

- *Rahmenbedingungen*
(Raum, Ausstattung, Lärm)
- *Verhalten des Dozenten*
(Aufbereitung und Darstellung des Stoffes, Leitungsfunktion)
- *Verhalten der Studierenden*
(Teilnahmeregelmäßigkeit, Aufmerksamkeit, Lerneinsatz)

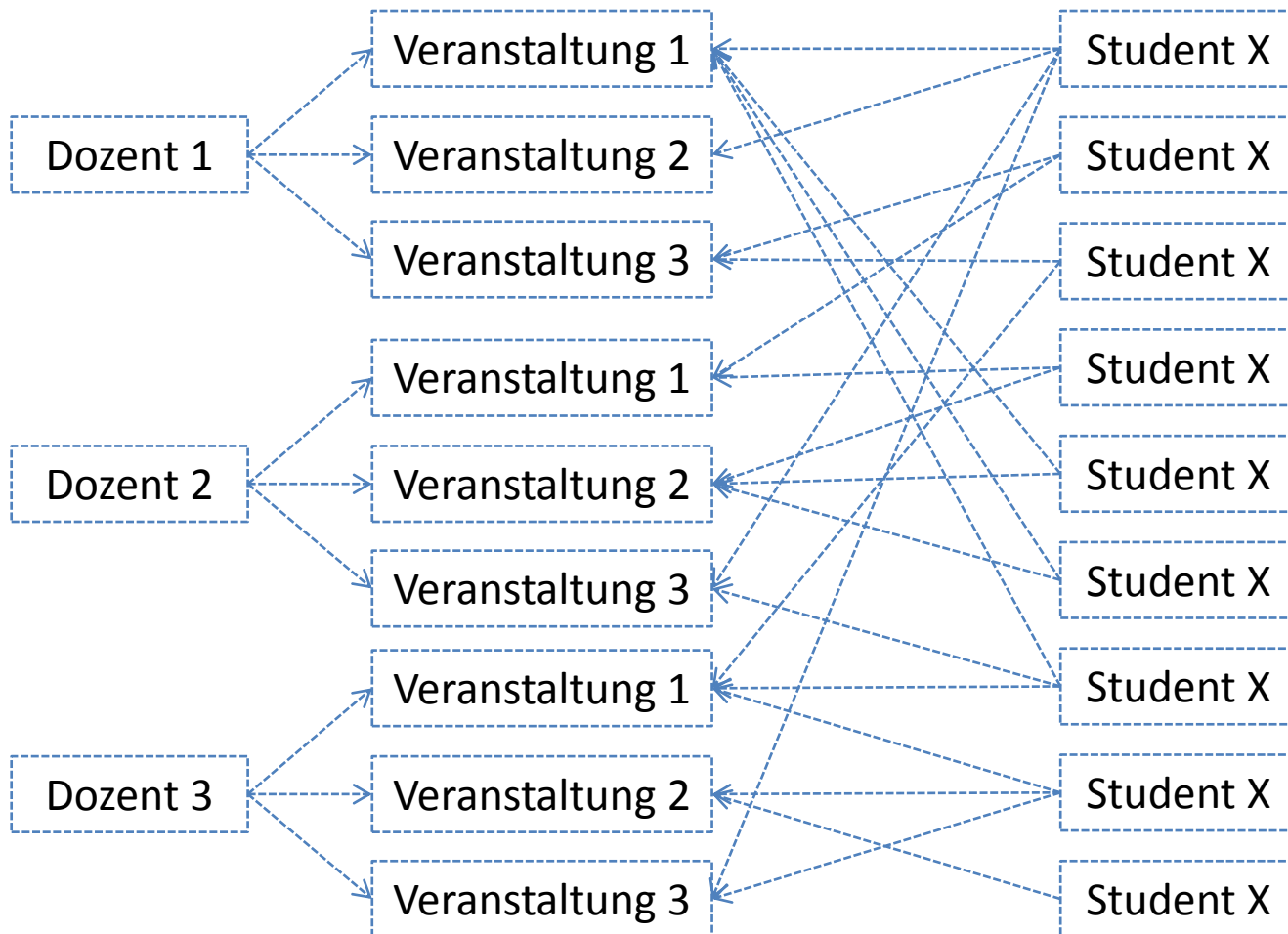
Ergebnisvariablen

(Veranstaltungsspezifisch, bilden Ergebnisse der Lehrveranstaltung ab)

- *Gesamteindruck*
(Attraktivität der Lehrveranstaltung)
- *Zuwachs an Wissen, Fertigkeiten, Kompetenzen*
(Fachwissen, praktische Anwendung, Fachübergreifendes Denken)

aus Born, Loßnitzer & Schmidt (2006)

1. Datensatzstruktur



- Komplexe Datenstruktur in vielen Lehrevaluationsdaten vorhanden
- Eindeutige Identifikation der Studierenden oft nicht möglich

1. Analysestrategie in der Lehrevaluation

- Analysen, Aussagen und Interpretationen auf Veranstaltungsebene (vgl. Yunker, 1983)
- Personen (itembasierte) Auswertung wird abgelehnt (vgl. Toland & De Ayala, 2005)
 - „Currently, researchers, such as Marsh and Roche, disallow any research that would use within-class data [...]” (Clayson, 2007)
- Auswertung erfolgt auf aggregierter Ebene (Aggregation je Item über Studierende innerhalb jeder Veranstaltung/ eines Dozenten)
- Die Verletzung der iid-Annahme: „independently and identically distributed“ (Skinner, Holt & Smith, 1989) bleibt oft unberücksichtigt.

1. Analysestrategie in der Lehrevaluation

Verletzung der iid-Annahme („independently and identically distributed“) erkennbar durch:

- Messwiederholung auf Dozentenebene und/oder Studentenebene
- Substantielle Design-Effekte > 2 (vgl. Muthén & Satorra, 1995)

Folgen:

- Verzerrung von Parametern, Standardfehlern und Fitstatistik (Du Toit & du Toit, 2008)

Lösung:

- Adjustierung der Standardfehler und Fitstatistik (Muthén & Satorra, 1995)
- Mehrebenen-Analyse

1. Lehrevaluation an der FSU Jena

2. Messmodell für PELVE

1. Modellfit bei Verletzung der iid-Annahme

2. Vergleich der Analysemethoden

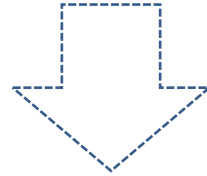
3. Zusammenfassung & Ausblick



2. Messmodell für PELVE

Born, Loßnitzer & Schmidt (2006):

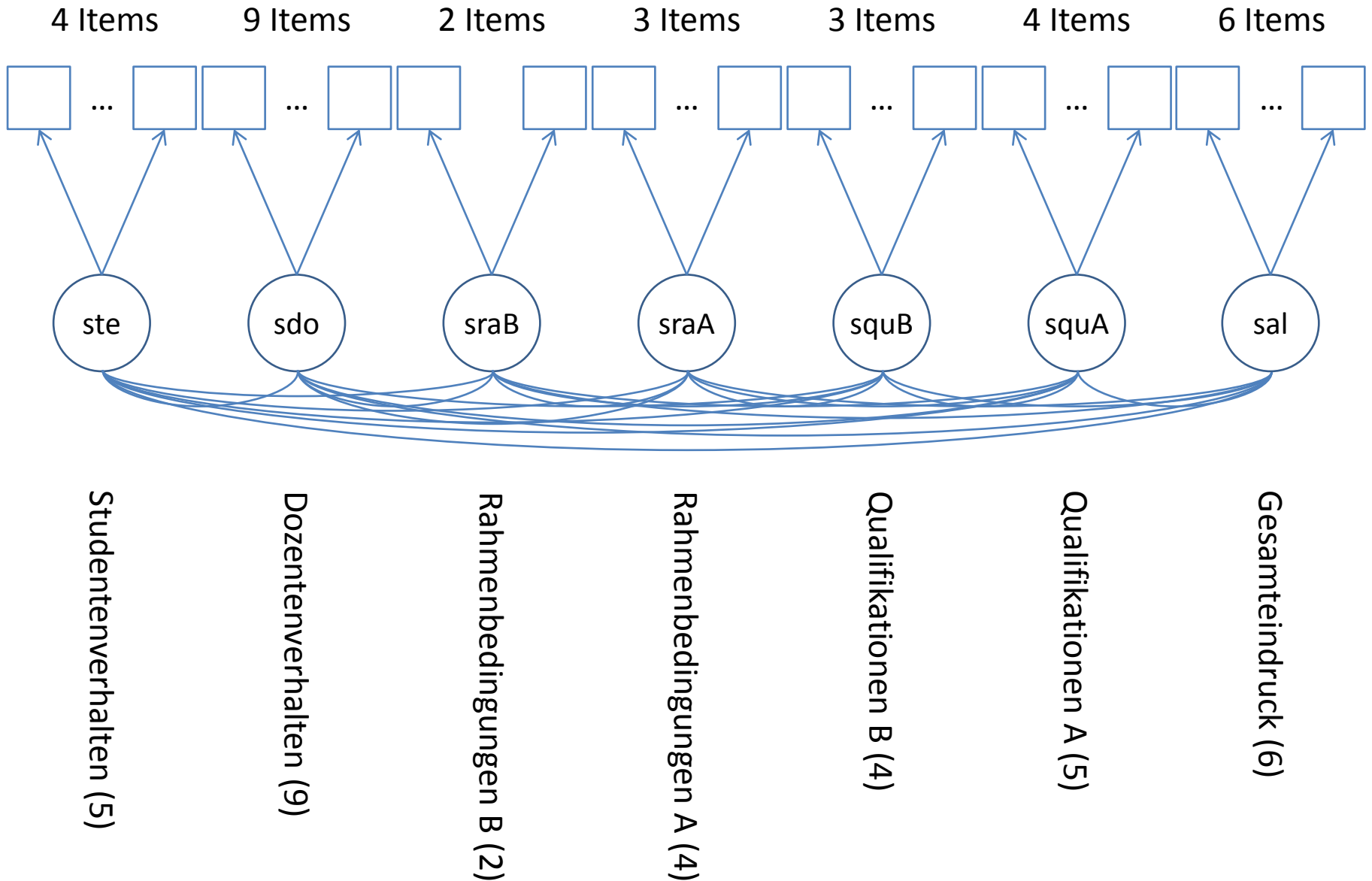
- 367 Seminare mit 2774 Studierenden über 3 Semester
- Seminarspezifische Items
- iid-Annahme verletzt
- Itembasierte Auswertung mit LISREL
- Theoretisch bestimmte Dimensionen werden abgebildet
 - 5 Faktoren erster Ordnung +1 für die Unterteilung des Kompetenzblocks



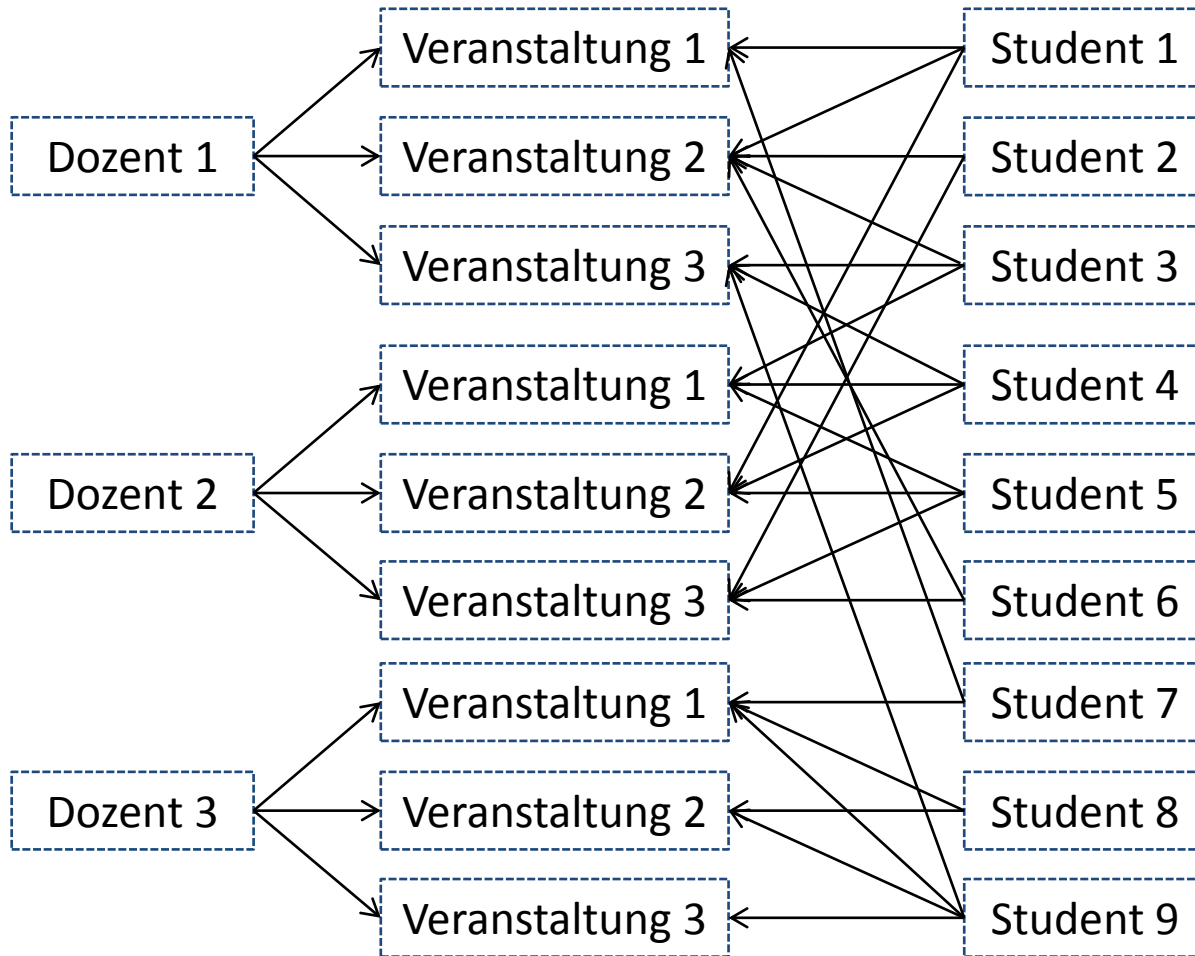
Implikationen für aktuelle Studie:

- Unterteilung der Rahmenbedingungen (7 Faktoren erster Ordnung)
- Sicherung der iid-Annahme
- Messmodell unabhängig vom Veranstaltungstyp
- Vergleich des Modellfits in Abhängigkeit der Datenbasis (aggregiert vs. itembasiert; iid Annahme)

2. Messmodell für PELVE



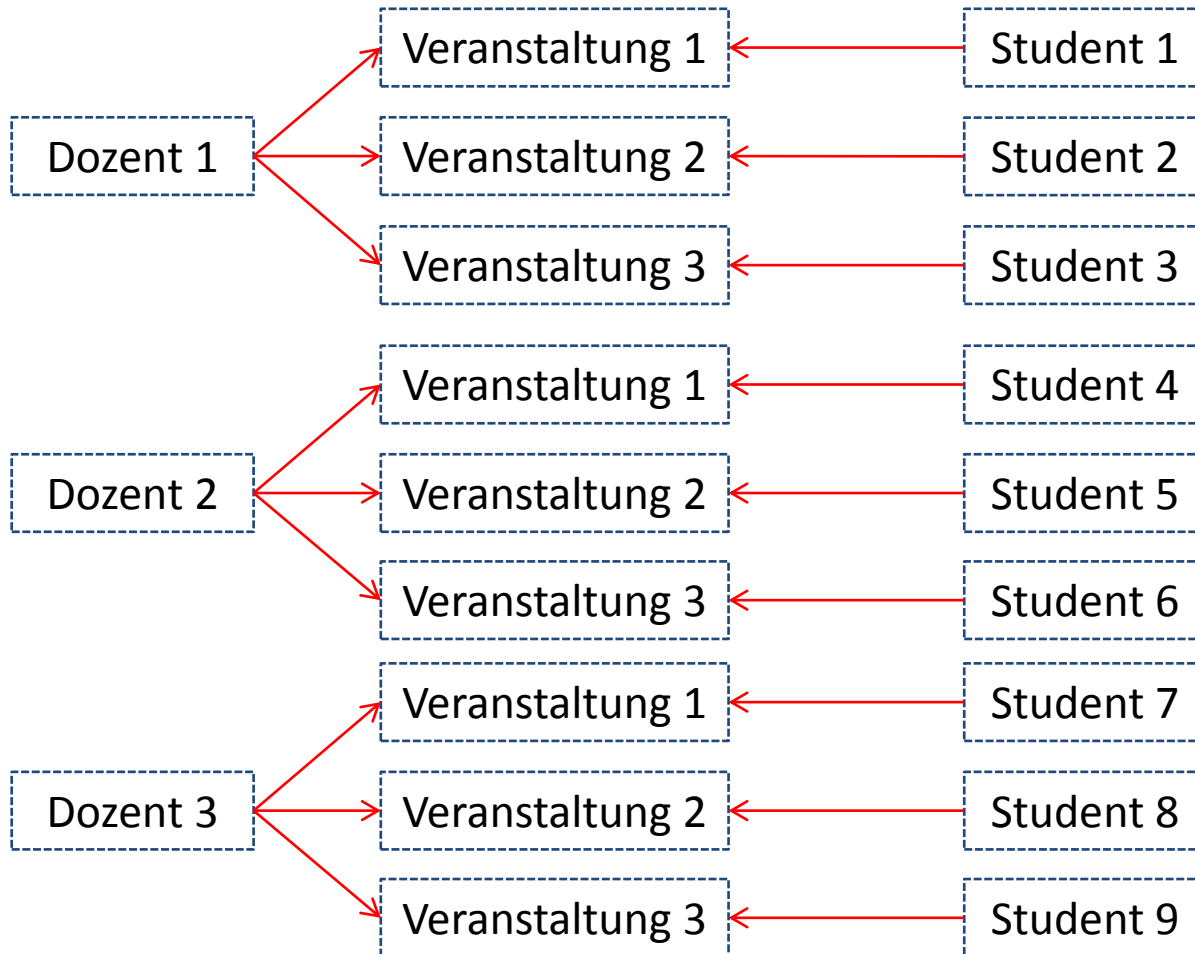
2. PELVE - Datensatz



Verschiedene Datenstrukturen möglich:

	Dozentenwiederholung	Studentenwiederholung
1	Ja	Ja
2	Ja	Nein
3	Nein	Ja
4	Nein	Nein

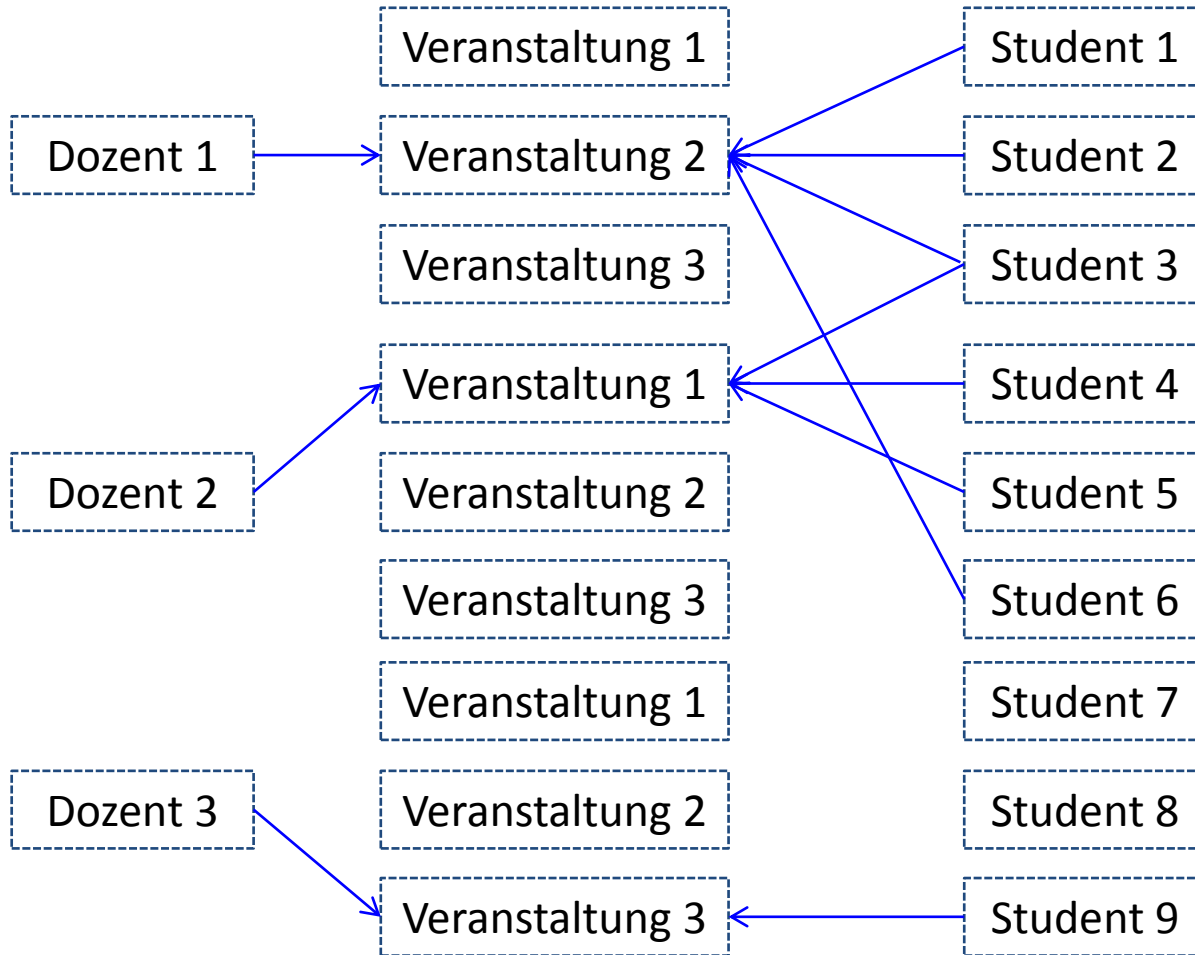
2. PELVE - Datensatz



Verschiedene Datenstrukturen möglich:

	Dozentenwiederholung	Studentenwiederholung
1	Ja	Ja
2	Ja	Nein
3	Nein	Ja
4	Nein	Nein

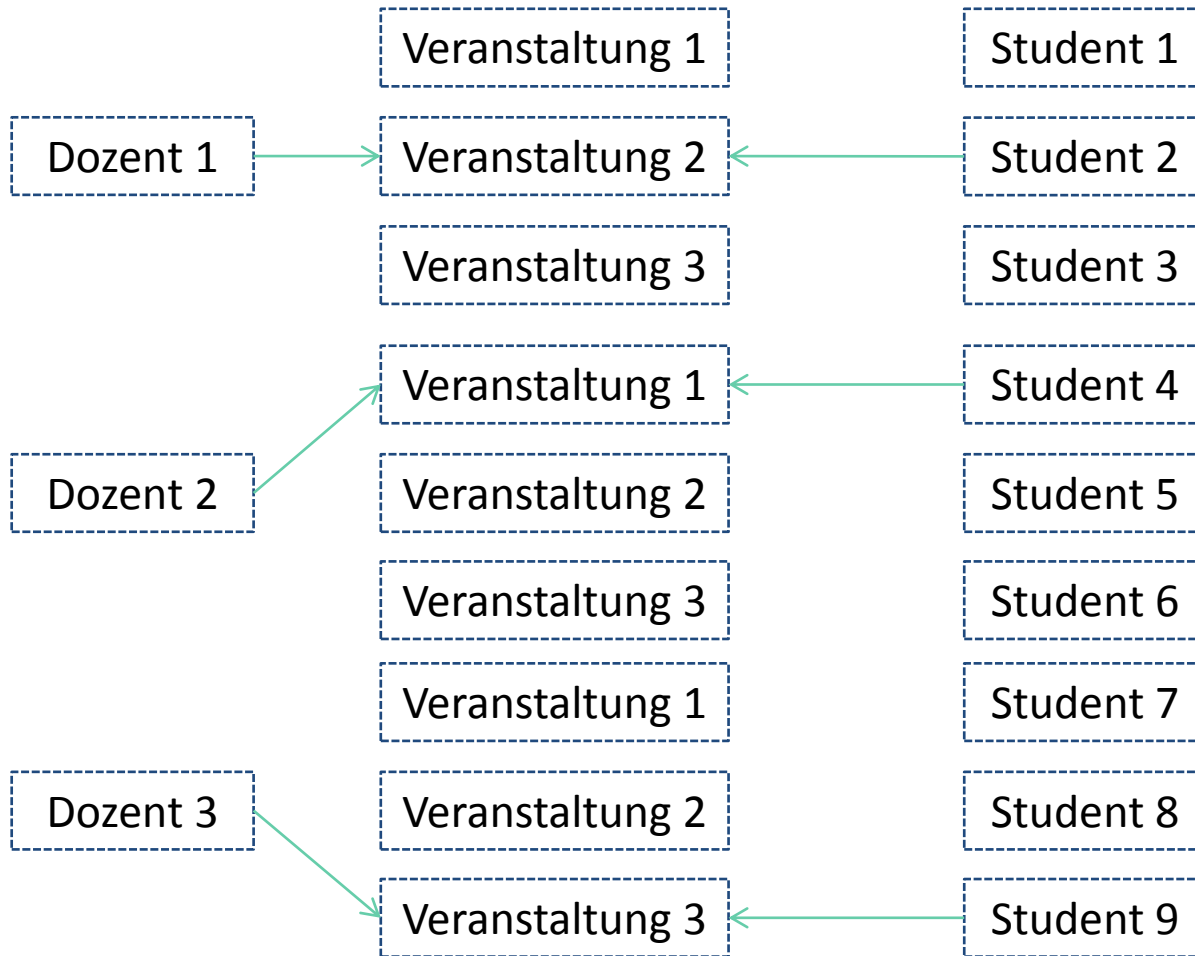
2. PELVE - Datensatz



Verschiedene Datenstrukturen möglich:

	Dozentenwiederholung	Studentenwiederholung
1	Ja	Ja
2	Ja	Nein
3	Nein	Ja
4	Nein	Nein

2. PELVE - Datensatz



Verschiedene Datenstrukturen möglich:

	Dozenten- wiederholung	Studenten- wiederholung
1	Ja	Ja
2	Ja	Nein
3	Nein	Ja
4	Nein	Nein

2. Datensätze

Gesamtdatensatz:

- 1 355 Dozenten
- 5 261 Veranstaltungen
- 129 042 Studierende

$N=200$ Datensätze aus dem Gesamtdatensatz für die Analysen

	Dozenten- wiederholung	Studenten- wiederholung	Studierende	Veranstaltungen
1	Ja	Ja	129042	5261
2	Ja	Nein	36459	5015.05 (11.21)
3	Nein	Ja	36249.32 (504)	1355
4	Nein	Nein	9443.63 (169)	1294.57 (6.08)

1,2 & 3: Verletzung der iid-Annahme durch „Messwiederholung“

4: Verletzung der iid-Annahme durch hohe Intraklassen-Korrelationen (*icc*) bzw. Design-Effekte (*deff*):

2. Fragestellungen

Ziel: Untersuchung der faktoriellen Struktur des PELVE Inventars der FSU Jena.

Frage 1: Welche Auswirkungen hat die Verletzung der iid-Annahme auf den Modell-Fit des PELVE-Messmodells?

Frage 2: Welche Analysemethode führt zur besten Passung des Modells?

1. Lehrevaluation an der FSU Jena
2. Messmodell für PELVE
 1. Modellfit bei Verletzung der iid-Annahme
 2. Vergleich der Analysemethoden
3. Zusammenfassung & Ausblick



2.1 Analysen

4 Analysen (A-D) für Frage 1: Welche Auswirkungen hat die Verletzung der iid-Annahme auf den Modell-Fit des PELVE-Messmodells?

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene		
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1 Ja / Ja	Analyse A	Analyse B	Analyse C	Analyse D
2 Ja / Nein				
3 Nein / Ja				
4 Nein / Nein				

2.1 Analysen

Analyse A: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der Analyse aggregierter Daten.

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene		
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1 Ja / Ja	Analyse A			
2 Ja / Nein				
3 Nein / Ja				
4 Nein / Nein				

2.1 Analysen

Analyse B-D: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der ...

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene		
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1 Ja / Ja				
2 Ja / Nein		Analyse B	Analyse C	Analyse D
3 Nein / Ja				
4 Nein / Nein				

- itembasierten Analyse (disaggregierter) Daten (**Analyse B**),
- unter Berücksichtigung der Erhebungsmethode (**Analyse C**)
- bzw. der Mehrebenenstruktur (**Analyse D**).

2.1 Analysen

1 Analyse für Frage 2: Welche Analysemethode führt zur besten Passung des Modells?

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene		
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1 Ja / Ja				
2 Ja / Nein				
3 Nein / Ja				
4 Nein / Nein	Analyse E			

- Inwiefern unterscheidet sich der Modellfit zwischen den Analysemethoden für den „bereinigten“ Datensatz?

2.1 Analyse A

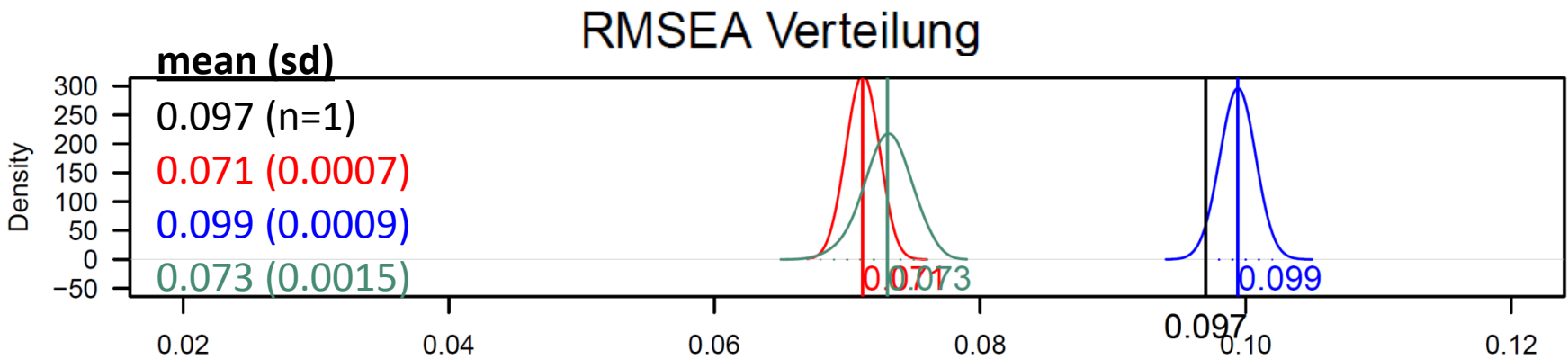
Analyse A: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der Analyse aggregierter Daten.

Dozentenwiederholung / Studentenwiederholung		Aggregierte Daten		Disaggregierte Daten		
		Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung		Rohdatenanalyse auf Itemebene		
		Kurse	Studierende	WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1	Ja / Ja	5261 (0)	129042 (0)			
2	Ja / Nein	5015.05 (11.21)	36459 (0)			
3	Nein / Ja	1355 (0)	36249.32 (504)			
4	Nein / Nein	1294.57 (6.08)	9443.63 (169)			

2.1 Analyse A: Modellfit für aggregierte Daten

Wie hängt die Modellgüte von der Datenstruktur ab?

- Aggregation der Ratings je Item innerhalb einer Veranstaltung



- Studentenwiederholung führt zur Überschätzung des RMSEA
- Bei Dozentenwiederholung wird der RMSEA im Vgl. zum iid-Datensatz etwas unterschätzt

- mit Dozentenwiederholung; mit Studentenwiederholung
- mit Dozentenwiederholung; ohne Studentenwiederholung
- ohne Dozentenwiederholung; mit Studentenwiederholung
- ohne Dozentenwiederholung; ohne Studentenwiederholung

2.1 Analyse B

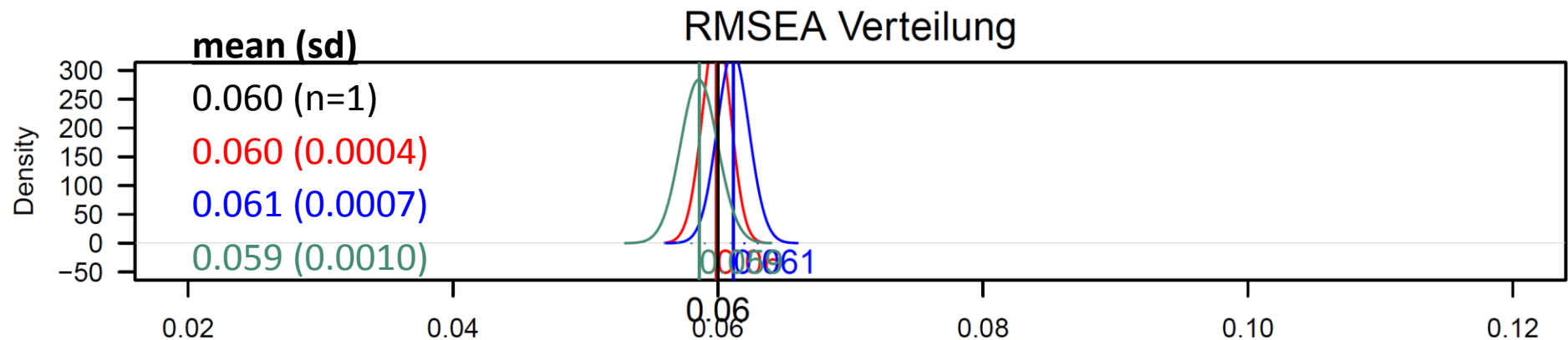
Analyse B: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der itembasierten Analyse (disaggregierter) Daten.

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene			
		WLSMV		WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
		Kurse	Studierende		
1 Ja / Ja		5261 (0)	129042 (0)		
2 Ja / Nein		5015.05 (11.21)	36459 (0)		
3 Nein / Ja		1355 (0)	36249.32 (504)		
4 Nein / Nein		1294.57 (6.08)	9443.63 (169)		

2.1 Analyse B: Modellfit für kategoriale Daten

Wie hängt die Modellgüte von der Datenstruktur ab?

- Kategoriale Daten je Item und Person liegen vor (WLSMV Schätzer)
- Hierarchische Struktur bleibt unberücksichtigt



- Sehr geringe Unterschiede in der Modellpassung trotz Verletzung der iid-Annahme

- mit Dozentenwiederholung; mit Studentenwiederholung
- mit Dozentenwiederholung; ohne Studentenwiederholung
- ohne Dozentenwiederholung; mit Studentenwiederholung
- ohne Dozentenwiederholung; ohne Studentenwiederholung

2.1 Analyse C

Analyse C: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der itembasierten Analyse (disaggregierter) Daten unter Berücksichtigung der Erhebungsmethode.

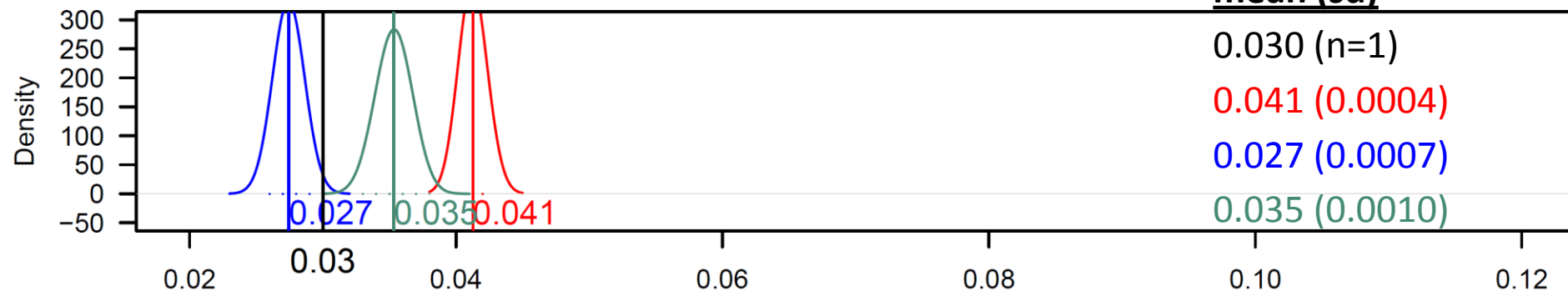
Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	WLSMV	Disaggregierte Daten Rohdatenanalyse auf Itemebene		WLSMV (Type = TWOLEVEL)
			WLSMV (Type=COMPLEX)		
			Kurse	Studierende	
1 Ja / Ja			5261 (0)	129042 (0)	
2 Ja / Nein			5015.05 (11.21)	36459 (0)	
3 Nein / Ja			1355 (0)	36249.32 (504)	
4 Nein / Nein			1294.57 (6.08)	9443.63 (169)	

2.1 Analyse C: Modellfit mit Type=COMPLEX

Wie hängt die Modellgüte von der Datenstruktur ab?

- Kategoriale Daten je Item und Person liegen vor (WLSMV Schätzer)
- Hierarchische Struktur wird berücksichtigt (TYPE = COMPLEX)

RMSEA Verteilung



- Unterschätzung des RMSEA wenn ein Studierender in mehrere Veranstaltungen eingeht
- Überschätzung des RMSEA wenn von einem Dozenten mehrere Veranstaltungen vorliegen
- Gute Modellpassung

- mit Dozentenwiederholung; mit Studentenwiederholung
- mit Dozentenwiederholung; ohne Studentenwiederholung
- ohne Dozentenwiederholung; mit Studentenwiederholung
- ohne Dozentenwiederholung; ohne Studentenwiederholung

2.1 Analyse D

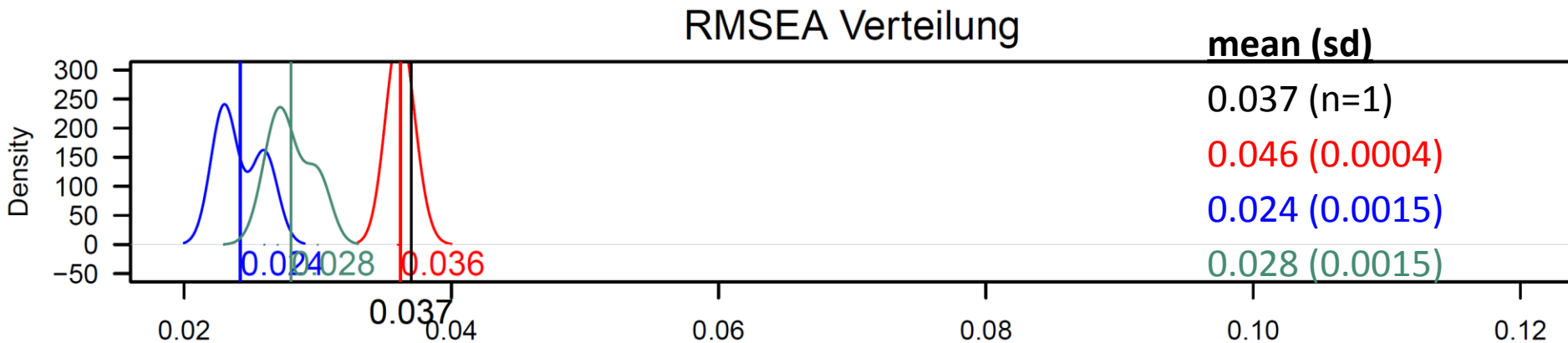
Analyse D: Welche Auswirkung hat die Verletzung der iid-Annahme auf den Modellfit bei der itembasierten Analyse (disaggregierter) Daten unter Berücksichtigung der Mehrebenenstruktur .

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene			
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)	
				Kurse	Studierende
1 Ja / Ja				5261 (0)	129042 (0)
2 Ja / Nein				5015.05 (11.21)	36459 (0)
3 Nein / Ja				1355 (0)	36249.32 (504)
4 Nein / Nein				1294.57 (6.08)	9443.63 (169)

2.1 Analyse D: Modellfit mit Type=TWOLEVEL

Wie hängt die Modellgüte von der Datenstruktur ab?

- Kategoriale Daten je Item und Person liegen vor (WLSMV Schätzer)
- Hierarchische Struktur wird berücksichtigt (TYPE = TWOLEVEL)



- Mehrebenenanalyse berücksichtigt die Datenstruktur
- Modellfit wird bei Studentenwiederholung überschätzt (zu kleiner RMSEA)

- mit Dozentenwiederholung; mit Studentenwiederholung
- mit Dozentenwiederholung; ohne Studentenwiederholung
- ohne Dozentenwiederholung; mit Studentenwiederholung
- ohne Dozentenwiederholung; ohne Studentenwiederholung

2.1 Ergebnisse: Abhängigkeit des Modellfits

1. Welche Auswirkungen hat die Verletzung der iid-Annahme auf den Modell-Fit des PELVE-Messmodells?
 - Stärkste Auswirkungen bei der Analyse aggregierter Daten
 - Bei der itembasierten Auswertung reagiert der Modellfit nur geringfügig auf die unterschiedlichen Datenstrukturen.
 - Unter Berücksichtigung der Erhebungsmethode wird die itembasierte Auswertung empfindlich gegenüber abhängigen (unberücksichtigten) Datenstrukturen.

1. Lehrevaluation an der FSU Jena
2. Messmodell für PELVE
 1. Modellfit bei Verletzung der iid-Annahme
 2. Vergleich der Analysemethoden
3. Zusammenfassung & Ausblick



2.2 Analyse E

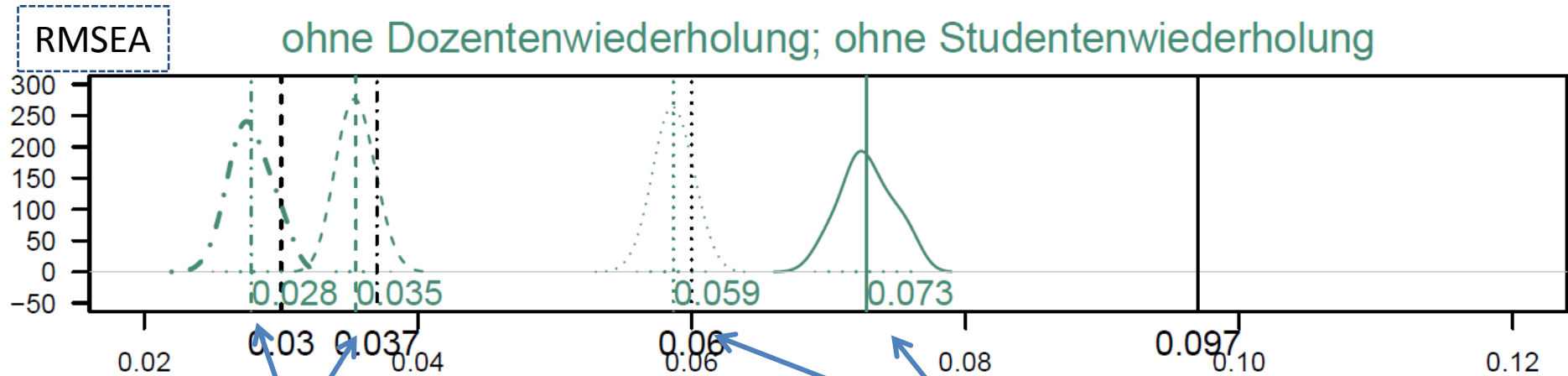
1 Analyse für Frage 2: Welche Analysemethode führt zur besten Passung des Modells?

Dozentenwiederholung / Studentenwiederholung	Aggregierte Daten Mittelwert jedes Items über alle Studierenden innerhalb einer Veranstaltung	Disaggregierte Daten Rohdatenanalyse auf Itemebene		
		WLSMV	WLSMV (Type=COMPLEX)	WLSMV (Type = TWOLEVEL)
1 Ja / Ja				
2 Ja / Nein				
3 Nein / Ja				
4 Nein / Nein	Analyse E			

- Inwiefern unterscheidet sich der Modellfit zwischen den Analysemethoden für den „bereinigten“ Datensatz?

2.2 Analyse E: Vergleich der Modellfits

ohne Dozentenwiederholung; ohne Studentenwiederholung



Berücksichtigung der Erhebungsmethode verbessert den Modellfit

Erhebungsmethode führt zur Verletzung der iid-Annahme und schlechterem Modellfit

— aggregierte Daten
- - - WLSMV & Type=COMPLEX
..... WLSMV
- - - MFA

4. Zusammenfassung & Ausblick

- Messmodell für aggregierte Daten reagiert empfindlich auf abhängige Datenstrukturen
- Effekte der Datenerhebung (Clustersampling) führen zur substantiellen Verletzung der iid-Annahme
- Korrektur / Berücksichtigung nötig (Type=COMPLEX; Mehrebenen-FA)

ToDo:

- Anpassung der Level-2-Faktorstruktur
- Interpretation Multilevel-SEM
- Vergleich der Personenparameterschätzungen
- Einfluss der Clustergröße

Danke



Anja Vetterlein
anja.vetterlein@uni-jena.de

Erik Sengewald
erik.sengewald@uni-jena.de