

## Mehrebenenanalyse: die wichtigsten 23 Modelle zur Analyse von dichotomen und metrischen Zwei- und Drei-Ebenen-Effekten (unter Verwendung von Mplus)

Urban, Dieter

Veröffentlichungsversion / Published Version

Forschungsbericht / research report

### Empfohlene Zitierung / Suggested Citation:

Urban, D. (2022). *Mehrebenenanalyse: die wichtigsten 23 Modelle zur Analyse von dichotomen und metrischen Zwei- und Drei-Ebenen-Effekten (unter Verwendung von Mplus)*. (2., überarb. Auflage) (Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart -SISS-, 51). Stuttgart: Universität Stuttgart, Fak. 10 Wirtschafts- und Sozialwissenschaften, Institut für Sozialwissenschaften. <https://doi.org/10.18419/opus-12146>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**gesis**  
Leibniz-Institut  
für Sozialwissenschaften

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der  
  
Leibniz-Gemeinschaft

**SISS:**  
**Schriftenreihe des Instituts für Sozialwissenschaften  
der Universität Stuttgart**

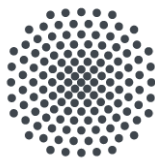
**No. 51 / 2022**

***Mehrebenenanalyse:***

***Die wichtigsten 23 Modelle  
zur Analyse von dichotomen und metrischen  
Zwei- und Drei-Ebenen-Effekten  
(unter Verwendung von Mplus)***

***2., überarbeitete Auflage***

***Dieter Urban***



**Universität Stuttgart**  
**Abteilung für Soziologie  
und empirische Sozialforschung**

**SISS:  
Schriftenreihe  
des Instituts für Sozialwissenschaften  
der Universität Stuttgart:  
No. 51 / 2022**

Mehrebenenanalyse:  
Die wichtigsten 23 Modelle  
zur Analyse von dichotomen und metrischen  
Zwei- und Drei-Ebenen-Effekten  
(unter Verwendung von Mplus)  
2., überarbeitete Auflage

.

Dieter Urban

**ISSN 09459197**

**Universität Stuttgart  
Institut für Sozialwissenschaften  
Abt. f. Soziologie mit dem Schwerpunkt sozialwissenschaftliche Forschungsmethoden  
70174 Stuttgart**

**Dieter Urban**

**Mehrebenenanalyse: Die wichtigsten 23 Modelle zur Schätzung von dichotomen und metrischen Zwei- und Drei-Ebenen-Effekten (unter Verwendung von Mplus)**

**2., überarbeitete Auflage**

SISS – Schriftenreihe des Instituts f. Sozialwissenschaften der Universität Stuttgart, No. 51/2022  
<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS51.pdf>

**Z U S A M M E N F A S S U N G:** Dieser Text liefert KEINE Einführung in die Methode der Mehrebenen-Analyse. Auch liefert dieser Text KEINE Einführung in die Datenanalyse mittels der Statistik-Software “Mplus”. Stattdessen setzt dieses Skript grundlegende Kenntnisse über die Methode der Mehrebenen-Analyse und über den Gebrauch der Statistik-Software “Mplus” voraus. Allerdings liefert der Text eine Auffrischung und Erweiterung dieser Kenntnisse.

Aber vor allem präsentiert das Skript insgesamt 23 ausgewählte Modelle zur Mehrebenen-Analyse. Diese Modelle werden beispielhaft für eine analytisch verstandene Logik und Methodik der Mehrebenen-Analyse vorgestellt (auch grafisch, wodurch sie viel leichter zu verstehen sind). Die 23 Modelle sind Muster-Modellierungen, die je nach Forschungsinteresse modifiziert (bzw. erweitert oder reduziert) werden können. Sie betreffen ausschließlich dichotome und metrische Zwei- und Drei-Ebenen- sowie Cross-Level-Effekte.

Die für die Analyse verwendeten Daten sind im Internet frei verfügbar. Alle Mplus-input files können mittels copy-and-paste aus dem Text übernommen werden.

**Dieter Urban**

**Multilevel Analysis: The 23 Most Important Models for Estimating Dichotomous and Metric Two- and Three-Level-Effects (By Using Mplus-Software)**

**2nd revised edition**

SISS – Schriftenreihe des Instituts f. Sozialwissenschaften der Universität Stuttgart, No. 51/2022  
<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS51.pdf>

**A B S T R A C T:** This script is NO introduction into the method of multilevel analysis. NOR does the text provide any introduction into using the software “Mplus” for statistical data analysis. Instead, this text presupposes basic knowledge of multilevel analysis and of applying “Mplus” to statistical analysis. However, this text provides also a refresher of this knowledge.

But above all, the script presents 23 well-selected models for multilevel analysis. These models are presented as examples of an analytically understood logic and methodology of multilevel analysis. For this purpose, they are also presented graphically, which makes them much easier to understand. The 23 models serve as exemplary analytic modellings that can be modified (extended or reduced) depending on research interests. They exclusively concern dichotomous and metric two-level- and three-level- as well as cross-level effects.

The data used for the analysis are freely available on the Internet. All Mplus-input files can be copied-and-pasted from the text of this script.

**Schlüsselwörter / key words** Mehrebenen-Analyse . Mehrebenenanalyse . MEA . Mehrebenenmodell . Mehrebenen-Modell . Regression . Mehrebenen-Regressionsanalyse . multilevel analysis . MLA . cross level effects . fixed effects . random effects . Mplus . Software

---

<b>0</b>	<b>Wozu dieses Skript ?</b> .....	<b>1</b>
<b>A</b>	<b>Warum Mehrebenen-Analyse?</b> .....	<b>3</b>
A1	Wann sollte eine Mehrebenenanalyse durchgeführt werden? (ICC, deff, HS-Tests).....	3
A2	Welche Voraussetzungen und Annahmen gelten für eine Mehrebenen-Analyse?.....	9
A3	Welche Vorteile hat eine Mehrebenen-Analyse?.....	13
<b>B</b>	<b>Einige ausgewählte Grundlagen der Mehrebenen-Analyse</b> .....	<b>17</b>
B1	Das Analysemodell .....	17
B2	Was bedeuten die Konfidenzintervalle von Random-Effekten .....	25
B3	Welche Bedeutung haben Kovarianzen zwischen den Fehlertermen? .....	26
B4	Wann und wie sollte zentriert werden? .....	29
B5	Wie sollten Modelle miteinander verglichen werden? (AIC, BIC, $R^2$ ) .....	31
<b>C</b>	<b>Mehrebenen-Analyse mit Mplus</b> .....	<b>35</b>
C1	Modellierung und Analysestrategie .....	35
C2	Mplus-Syntax und Mehrebenen-Analyse .....	38
C3	Standardisierung .....	41
<b>D</b>	<b>Mehrebenen-Regressionsanalyse für 2 Ebenen</b> .....	<b>43</b>
D1	Die Gliederung .....	43
D2	Der Datensatz .....	44
D3	Tabellarische Klassifikation von 14 Zwei-Ebenen-Modelle .....	45
D4	Systematische Erläuterung von 14 Zwei-Ebenen-Modellen .....	47
	(Modell 01 bis Modell 14)	
<b>E</b>	<b>Mehrebenen-Regressionsanalyse für 3 Ebenen</b> .....	<b>149</b>
E1	Die Gliederung .....	149
E2	Der Datensatz .....	150
E3	Tabellarische Klassifikation von 9 Drei-Ebenen-Modellen .....	151
E4	Systematische Erläuterung von 9 Drei-Ebenen-Modellen .....	153
	(Modell 20 - Modell 28)	
<b>F</b>	<b>Tabellarischer Modellvergleich</b> .....	<b>223</b>
<b>G</b>	<b>Gebräuchliche Notation von Komponenten der Mehrebenen-Analyse</b> .....	<b>224</b>
<b>H</b>	<b>Literatur</b> .....	<b>227</b>

## 0 Wozu dieses Skript ?

Das vorliegende Skript

- ist KEINE Einführung in die Methode der Mehrebenen-Analyse,
- ist KEINE Einführung in die Datenanalyse mittels der Statistik-Software “Mplus”.

Vielmehr setzt das vorliegende Skript voraus:

- grundlegende Kenntnisse über die Methode der Mehrebenen-Analyse,
- grundlegende Kenntnisse zum Gebrauch der Statistik-Software “Mplus”.

Allerdings liefert das vorliegende Skript:

- eine Auffrischung von Kenntnissen über die Logik und Methodik der Mehrebenen-Analyse,
- eine Erweiterung von Kenntnissen über den Einsatz von Mplus in der Mehrebenen-Analyse.

Das vorliegende Skript zeichnet sich vor allem dadurch aus,

- dass es 23 Modelle zur Mehrebenen-Analyse mit Mplus beispielhaft präsentiert, sodass Anwender/innen nach den Beispielen suchen können, die sie für ihre Analysezwecke komplett oder angepasst übernehmen können.

Das vorliegende Skript folgt einer speziellen Auffassung von Methodendidaktik insofern als:

- alle Modelle der Mehreben-Analyse beispielhaft vorgestellt werden,
- alle Modelle komplett vorgestellt werden (einschließlich ihrer Logik, ihrer Formalismen, ihrer bildhaften Darstellung, ihrer Interpretation und der dazugehörigen Mplus-Input- und Output-Files),
- sodass Anwender/innen alle angebotenen Informationen sowohl “1 zu 1” als auch “modifiziert” für ihre Zwecke übernehmen können.

Hinsichtlich der Methodik der hier vorgestellten Modelle erfolgt in diesem Skript eine Einschränkung dadurch,

- dass nur Zwei- und Drei-Ebenen-Modelle erläutert werden,
- dass nur Modelle mit dichotomen und metrischen Variablen beschrieben werden,
- dass alle Modell mit der robusten Maximum-Likelihood-Methode von Mplus (MLR) geschätzt werden.

Mit “Copy-and-Paste” können alle im Skript abgedruckten Mplus-Inputfiles

- direkt in Mplus eingelesen werden,
- (es müssen nur in den Mplus-Inputfiles die speziellen Angaben zum jeweils zu analysierenden Datensatz angepasst werden).

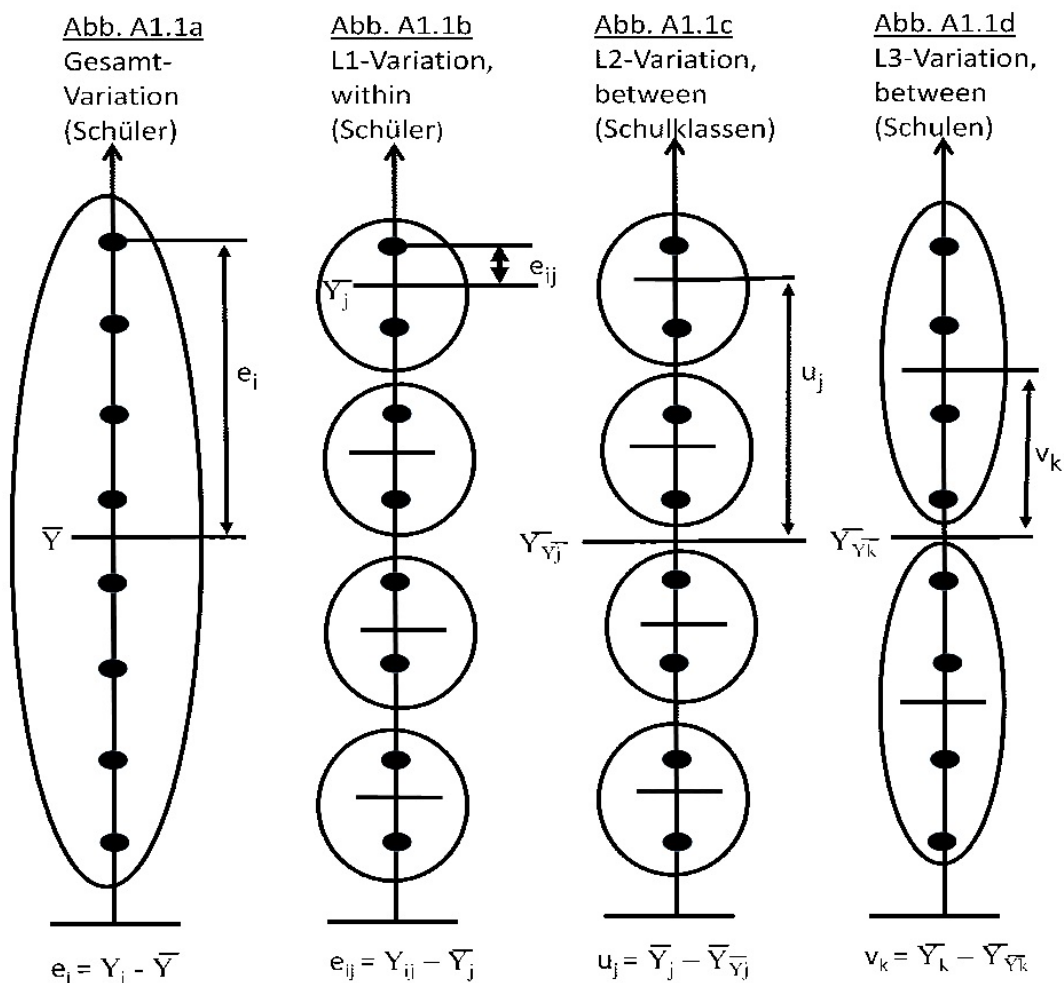
## A Warum Mehrebenen-Analyse?

### A1 Wann sollte eine Mehrebenen-Analyse durchgeführt werden?

Um die Methodik der Mehrebenen-Analyse richtig verstehen zu können, sollte der Anwender<sup>1</sup> mit dem Modell der statistischen Varianzzerlegung vertraut sein. Wir wollen deshalb im Folgenden zunächst das Modell der Varianzzerlegung an einem stilisierten Beispiel verdeutlichen und werden dann zu unserer eigentlichen Frage zurückkehren: "Wann sollte eine Mehrebenen-Analyse durchgeführt werden?"

In unserem stilisierten Beispiel gehen wir davon aus, dass wir in einem Forschungsprojekt eine große Anzahl von Messwerten über Schüler:innen (Analyseebene L1), über deren Schulklassen (Analyseebene L2) und über deren Schulen (Analyseebene L3) besitzen (der diesbezügliche Datensatz wird in Kap. E2 ausführlich erläutert).

Abb. A1.1: Stilisierte Darstellung des Prinzips der Varianzzerlegung



<sup>1</sup> Das hier und im Folgenden benutzte generische Maskulinum beschreibt die Verwendung der männlichen grammatikalischen Form unabhängig vom Geschlecht der beschriebenen Person.

Zunächst soll uns nur die Mathematikleistung der Schüler:innen interessieren. Deren Mathematikleistung wurde nach einem Punktesystem gemessen und in der Variablen MATH abgespeichert. Die Variable MATH sei unsere zu erklärende Variable Y. Für unserer Erläuterung der Varianzzerlegung gehen wir davon aus, dass uns nur die Y-Werte von acht Schüler:innen zur Verfügung stehen.

Die Abbildung A1.1 verdeutlicht die Logik der Varianzzerlegung. In den Abbildungen A1.1a bis A1.1d wird die Zerlegung der Gesamtvariation von Y in die drei Teilvariationen: Within-Variation, Between-Variation (L2) und Between-Variation (L3) gezeigt. Die vier Abbildungen verdeutlichen die jeweiligen Variationen von Y. Aus diesen Variationen können auch leicht die entsprechenden Varianzen berechnet werden, was hier aber nicht geschehen soll, weil dadurch die Abbildungen leicht unanschaulich würden.<sup>2</sup>

Die Abbildung A1.1a verdeutlicht die Gesamtvariation von Y. Die Gesamtvariation ergibt sich aus den Abständen von acht Personen (bzw. von deren Y-Werten) zu ihrem gemeinsamen Mittelwert. Vier der Datenpunkte liegen oberhalb und weitere vier Datenpunkte liegen unterhalb der Mean-Linie. Je größer die Entfernung der Datenpunkte von der Mean-Linie ist, umso größer wird auch die Gesamtvariation werden. Mit “ $e_i$ ” wird der Abstand einer jeden Person zum Gesamtmittelwert bezeichnet. In Abbildung A1.1a wird beispielhaft für einen Schüler “i” der Abstand von dessen individuellem MATH-Wert zu dem mittleren MATH-Wert durch eine senkrechte Linie mit der Länge “ $e_i$ ” markiert.

Die Abbildung A1.1b zeigt, wie der Teil der Variation entsteht, der als Within-Variation bezeichnet wird und der ein Teil der Gesamtvariation von Y ist. Dazu werden jeweils alle Y-Werte derjenigen Schüler:innen betrachtet, die sich in jedem Kontext/Cluster befinden (hier: in einer jeden Schulklasse). In unserer Abbildung sind in jedem Cluster bzw. in jeder Schulklasse nur zwei Schüler:innen vertreten. Für jedes dieser Cluster wird ein clusterspezifischer Y-Mittelwert ermittelt (hier: der MATH-Mittelwert von jeweils zwei Schüler:innen) und sodann der Abstand zwischen den einzelnen Y-Werten der Schüler:innen ( $Y_{ij}$ ) und dem jeweiligen clusterspezifischen Y-Mean berechnet. Die Summe dieser Abstände (genannt: “ $e_{ij}$ ”) ergibt die Within-Variation von Y. In Abbildung A1.1b wird dieser Abstand durch eine kleine senkrechte Linie, die rechts von der Y-Achse verläuft, gekennzeichnet.

In Abbildung A1.1c wird die Berechnung der Between-Variation verdeutlicht. Relevant für die Variationsberechnung sind jetzt nur noch die Abstände (genannt: “u”) zwischen jedem Cluster-Mittelwert (bzw. jedem Y-Mittelwert in jeder Schulklasse) und dem Grand-Mittelwert (dem Mittelwert aller kontextspezifischen Mittelwerte). Der Abstand für die j-te Schulklasse wird in der Abbildung A1.1c mit einer senkrechten Linie rechts von der Y-Achse veranschaulicht. Die Summe dieser Abstände (genannt: “ $u_j$ ”) ergibt die Between-Variation von Y.

In Abbildung A1.1d wird ebenfalls eine Between-Variation von Y veranschaulicht, allerdings diesmal nur für die beiden L3-Cluster/Kontexte (hier: Schulen). Relevant für die Variation sind wiederum nur die Abstände (genannt “ $v_k$ ”) zwischen zwei Mittelwerten, nämlich zwischen jedem L3-Cluster-Mittelwert (Y-Mittelwert einer jeden Schule) und dem Grand Mean aller L3-Kontext-Mittelwerte. Ein solcher Abstand wird in der Abbildung als senkrechte Linie

---

<sup>2</sup>  $\text{Var}(e) = 1/n \sum(e^2)$



rechts von der Y-Achse veranschaulicht. Die Summe dieser Abstände (genannt: “ $v_k$ ”) ergibt eine weitere Between-Variation von Y.

Nachdem wir das Modell der Varianzzerlegung (bzw. korrekter: der Variationszerlegung) verdeutlicht haben, kehren wir nun zurück zu unserer eigentlichen Ausgangsfrage: “Wann sollte eine Mehrebenen-Analyse durchgeführt werden?”

Eine Mehrebenen-Analyse sollte dann durchgeführt werden, wenn zumindest die folgenden drei Kriterien erfüllt sind (weitere Kriterien werden in Kap. A2 benannt):

- (1) L1-Messwerte müssen verschiedenen Clustern zuzuordnen sein (z.B. schülerbezogene Messwerte müssen jeweils bestimmten Schulklassen zuzuordnen sein) und die Cluster müssen einen prägenden Einfluss auf die L1-Messwerte ausüben, sodass die Messwerte ihre Unabhängigkeit verlieren (z.B. kann die Zugehörigkeit zu einer bestimmten Schulklasse mit einem best. Leistungsniveau in Mathematik das individuelle Leistungsvermögen eines jeden Schülers beeinflussen, sodass die individuellen Leistungswerte ihre Unabhängigkeit voneinander verlieren).
- (2) Es müssen L1-Messwerte und ggf. auch L2-Messwerte über mindestens zwei hierarchisch geschachtelte (nested) Analyseebenen (L1 und L2) vorliegen (z.B. über Schüler und über Schulklassen).
- (2) Der Anteil der “Between-Varianz” sollte mindestens 10% von der Gesamtvarianz der abh. Variablen (z.B. von der “Lebenszufriedenheit” befragter Schüler) betragen. Wenn es keine zu erklärende Varianz auf der zweiten (oder noch höheren) Ebene des Mehrebenen-Designs gäbe, wäre es sinnlos, eine MLA durchzuführen.

Die 10%-Daumenregel gilt nach einer unter Forschenden weit verbreiteten Konvention.<sup>3</sup> Sie geht davon aus, dass bei einer Datenstruktur, wie sie oben unter Pkt. 1 beschrieben wird, die Gesamtvarianz der abh. L1-Variablen in eine Within- und eine Between-Teilvarianz zerlegt werden kann. In welcher Weise dies möglich ist und welche Überlegungen damit verbunden sind, haben wir mit Abbildung A1.1 erläutert.

Abzulesen ist der relative Anteil der Between-Teilvarianz an der Größe des “Intra-Class-Correlation-Koeffizient” (ICC-Koeffizient bzw. Intra-Klassen-Korrelationskoeffizient). Der ICC berichtet darüber, wie hoch der prozentuale Varianzanteil ist, der durch Unterschiede zwischen den Clustern/Kontexten entsteht. Ist der ICC hoch genug, so sollte zur Analyse von Abhängig-

---

<sup>3</sup> Es gibt auch Forscher, die einen noch kleineren Varianzanteil als die hier genannten 10% als ausreichend für die Durchführung einer sinnvollen MLA betrachten. So wird u.a. ein Grenzwert von 5% oder noch kleiner vorgeschlagen (vgl. Huang 2018: 2f). Wir wollen hier allerdings an einem 10% Richtwert festhalten, der jedoch in begründeten Einzelfällen durchaus reduziert werden kann.

keitsstrukturen keine OLS-Regression, sondern eine MLR/MLA<sup>4</sup> durchgeführt werden. Ziel der MLA ist es dann, durch Einführung von L2-Variablen (und cross-level Interaktionen) die Between-Varianz auf null zu bringen.

Wir wollen im Folgenden den ICC-Koeffizienten für 2- und 3-Ebenen-Modelle kurz erläutern:

a) für 2-Ebenen-Modelle:

Das Intercept-Only-Modell für 2-Ebenen-Modelle ist:

(vgl. dazu auch die Ausführungen zu Modell 01 in Kap. D4)

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

Der ICC-Koeffizient (für 2 Ebenen) ist:

$$\rho = \text{Var}(u_{0j}) / [ \text{Var}(u_{0j}) + \text{Var}(e_{ij}) ]$$

$\rho$  = griechischer Buchstabe "rho"

oder:

$$\rho = \text{L2-Varianz} / [\text{L2-Varianz} + \text{L1-Varianz}]$$

Der ICC-Koeffizient informiert mithin über den Anteil der L2-Varianz "Var(u)" an der Gesamtvarianz "Var(u)+Var(e)" und sollte für eine MLA mindestens einen Wert von 0,10 (10%) aufweisen, d.h. der Anteil der L2-Varianz an der Gesamtvarianz sollte möglichst größer sein als 10%, damit es sinnvoll ist, eine MLA durchzuführen.<sup>5</sup> In unserem Beispielsmodell 01 (vgl. Kap. D4) beträgt der ICC-Koeffizient "0,185". Somit ist die Clusteraufteilung der Daten für 18,5% der Gesamtvarianz in unserem Modell verantwortlich. Insgesamt 18,5% der Varianz in der abh. Variablen (hier: "Lebenszufriedenheit") ist abhängig von Differenzen zwischen Einheiten auf der L2-Ebene (hier: "Länder") und es sollte deshalb eine MLA durchgeführt werden, um Prädiktoren zu identifizieren, die weitere Anteile der L2-Varianz binden können. Die Modellschätzung 01 (in Kap. D4) zeigt, wie der ICC-Koeffizient direkt aus dem betreffenden Mplus-Programm abgerufen werden kann:

Mplus: Variable LEZU / Intraclass Correlation = 0,185

b) für 3-Ebenen-Modelle:

Das Intercept-Only-Modell für 3-Ebenen-Modelle ist (vgl. dazu auch die Ausführungen zu Modell 20):

$$Y_{ijk} = \gamma_{000} + u_{0jk} + v_{00k} + e_{ijk}$$

Der ICC-Koeffizient, der den Anteil der Varianz auf der Ebene der Schulklassen (L2) bestimmt, ist:

---

<sup>4</sup> MLR = Maximum Likelihood-Schätzung, robust (vgl. Kap. A2, Pkt. 4) / MLA = Mehrebenen-Analyse.

<sup>5</sup> Mehr dazu in Fußnote 3.

$$\rho = \text{Var}(u_{0jk}) / [ \text{Var}(u_{0jk}) + \text{Var}(v_{00k}) + \text{Var}(e_{ij}) ] \quad \rho = \text{griechischer Buchstabe "rho"}$$

Der ICC-Koeffizient, der den Anteil der Varianz auf der Ebene der Schulen (L3) bestimmt, ist:

$$\rho = \text{Var}(v_{00k}) / [ \text{Var}(u_{0jk}) + \text{Var}(v_{00k}) + \text{Var}(e_{ij}) ] \quad \rho = \text{griechischer Buchstabe "rho"}$$

Die beiden ICC-Koeffizienten geben an, wie groß der prozentuale Anteil der L2- sowie der L3-Varianz an der Gesamtvarianz der Variablen ist. Nach der bereits oben erwähnten Daumenregel der MLA sollte der Anteil mindestens 10% betragen, damit eine MLA deutliche Vorteile gegenüber einer klassischen Ein-Ebenen-Analyse aufweist.<sup>6</sup> Die Modellschätzung 20 (in Kap. E4) zeigt, wie die beiden ICC-Koeffizienten direkt aus dem betreffenden Mplus-Programm heraus abgerufen werden können:

Mplus: SUMMARY OF DATA / Estimated Intraclass Correlations for the Y Variables for TEACHID level / Variable MATH: 0,082

Mplus: SUMMARY OF DATA / Estimated Intraclass Correlations for the Y Variables for SCHCODE level / Variable MATH: 0,124

Auch der sogenannte "Designeffekt" (deff) kann benutzt werden, um zu ermitteln, ob es zur Durchführung einer Datenanalyse sinnvoll ist, eine MLA durchzuführen. Der Designeffekt liefert ein Maß zur Beurteilung der Adäquanz einer Mehrebenen-Analyse. Der Effekt berechnet sich nach der folgenden Formel:

$$\text{deff} = 1 + [ (\text{average cluster size}) - 1 ] * \text{ICC}$$

Nach dieser Formel gilt: Je größer die L2-Samplegröße ist (bzw. je mehr Beobachtungen in einem Cluster sind), und je größer der Anteil der Varianz zwischen den Makroebenen an der Gesamtvarianz von Y ist, umso größer wird der Designeffekt für die entsprechende L-Ebene sein. Der Designeffekt für L2 unseres Modell 20 beträgt somit:<sup>7</sup>

$$\text{deff} = 1 + [ (11,47) - 1 ] * 0,082 = 1,85$$

Als Richtwert für einen Designeffekt, der eine MLA legitimieren kann, hat sich in der Forschung ein Wert von 2,00 durchgesetzt. Ein Designeffekt, der eine MLA rechtfertigen kann, sollte möglichst größer als 2,00 sein. Dementsprechend wäre es zur Analyse der im Datensatz "ch4three.DAT" (vgl. Kap. E2) enthaltenen Daten eher nicht gerechtfertigt, eine MLA durchzuführen.

Anders verhält es sich mit dem Designeffekt des gleichen Datensatzes für L3:<sup>8</sup>

---

<sup>6</sup> Mehr dazu in Fußnote 3.

<sup>7</sup> Vgl. dazu Kap. E2 sowie die Erläuterungen zu Modell 20 in Kap. E4.

<sup>8</sup> Vgl. dazu Kap. E2 sowie die Erläuterung von Modell 20 in Kap. E4.

$$\text{deff} = 1 + [(45,52) - 1] * 0,124 = 6,52$$

Die Größe des L3-Designeffekts von 6,52 signalisiert, dass es sehr wohl sinnvoll sein kann, die Strukturen des Datensatzes "ch4three.DAT" mit einer MLA auszuwerten.

Ob Datenanalysen mittels Schätzung von Mehrebenen-Modellen durchgeführt werden sollten, lässt sich auch anhand von Heteroskedastizitäts-Tests (HS-Tests) ermitteln. Denn unberücksichtigte Kontexteffekte führen oftmals dazu, dass sich die Varianzen von OLS-Residuen ungleich verteilen (vgl. dazu Urban/Mayerl 2018: 252-282), sodass an den Streuungsungleichheiten der Residuen einer OLS-Schätzung zu erkennen ist, ob die untersuchten Daten gegen die Unabhängigkeitsvoraussetzung statistischer Datenanalyse verstoßen. Lässt sich ein solcher Verstoß erkennen, dann wäre es sicherlich sinnvoll, Mehrebenen-Modelle zur Datenanalyse einzusetzen.

Von den vielen Möglichkeiten zum Test auf Heteroskedastizität möchten wir hier insbesondere den Glejser-Test empfehlen (Urban/Mayerl 2018: 266f). Er lässt sich einfach durchführen, ist intuitiv verständlich und erbringt in aller Regel eindeutig zu interpretierende Ergebnisse.

Eine Voraussetzung zur Durchführung des Glejser-Tests ist, dass die Heteroskedastizität (bzw. die Varianzungleichheit der Residuen einer OLS-Schätzung) eine Funktion eines X-Prädiktors oder eine Funktion von mehreren X-Prädiktoren ist. Die Residuen der Linearschätzung von Y durch X müssen dann zentriert, in absolute Werte transformiert und als eigenständige Variable abgespeichert werden.<sup>9</sup> Die Varianz der transformierten Residuen kann als abhängige Variable in einer erneuten OLS-Schätzung auf Effekte der X-Prädiktoren zurückgeführt werden. Ist zumindest einer der Regressionskoeffizienten in der zweiten OLS-Schätzung signifikant, ist dies ein Hinweis auf Heteroskedastizität und damit auch ein Hinweis auf die Nicht-Unabhängigkeit beobachteter Daten, die durch Kontexteffekte entstanden sein kann.

---

<sup>9</sup> Dies ist notwendig, damit positive und negative Residuen sich nicht gegenseitig neutralisieren.

## A2 Welche Voraussetzungen und Annahmen gelten für eine Mehrebenen-Analyse?

Im Folgenden wollen wir die wichtigsten Voraussetzungen und Annahmen auflisten, die für die Durchführung einer sinnvollen Mehrebenen-Analyse (MLA) einzuhalten sind. Dabei konzentrieren wir uns auf den Typus von MLA, der im vorliegenden Skript erläutert und von uns zur Modellschätzung eingesetzt wird: eine Mehrebenen-Regressionsanalyse mit einem robusten ML-Schätzverfahren (MLR) zur Ermittlung von dichotomen und metrischen Zwei- sowie Drei-Ebenen- und cross-level-Effekten.

### (1) lineare Modellspezifikation

Für eine Mehrebenen-Analyse sollte ein Linearmodell (d.h. ein Modell mit ausschließlich linearen Beziehungen zwischen den Modellvariablen inkl. nicht-linearer Effekte, die intrinsisch linear sind<sup>10</sup>) mittels analytischer oder theoretischer Argumentation begründet werden. Das Modell muss eine abhängige Variable "Y" und einen oder mehrere Prädiktoren "X, Z, W"<sup>11</sup> auf zwei oder drei hierarchisch geschachtelten (nested) Analyseebenen "L1, L2, L3" aufweisen und sollte keine Spezifikationsfehler enthalten (z.B. dadurch, dass wichtige Kontextvariablen fehlen, vgl. dazu Kap. A3). Zudem sollte spezifiziert werden, welche der linearen Effekte auf welcher Ebene "fixed" und/oder "random" und/oder "cross-level"- Effekte sind.

### (2) hinreichende Varianz von abh. und unabh. Variablen

Nur wenn die abhängige Y-Variable eine ausreichende Varianz aufweist, können Teile dieser Varianz durch Effekte der unabhängigen X/Z/W-Variablen gebunden bzw. statistisch erklärt werden. Und nur wenn die unabhängigen Variablen eine ausreichende Varianz aufweisen, können diese Variablen auch eingesetzt werden, um Varianzanteile der abhängigen Variablen zu binden bzw. statistisch zu erklären. Wie groß die zu erklärende Varianz und wie groß die erklärende Varianz minimal sein müssen, lässt sich leider nicht allgemein festlegen, sondern ergibt sich jeweils aus der Komplexität des zu untersuchenden Modells.

### (3) metrisches Skalenniveau/multivariate Normalverteilung

Die Maximum-Likelihood-Schätzung (s. Pkt. 4) geht von einer multivariaten Normalverteilung aus, die der beobachteten Kovarianzmatrix der Sampledaten zugrunde liegt und die dafür sorgt, dass im Schätzverfahren optimale Schätzwerte für die Populationsparameter erzielt werden können. Deshalb sollte die Y-Variable kontinuierlich bzw. metrisch skaliert und annäherungsweise normalverteilt sein. Auch die X/Z/W-Variablen sollten ebenfalls kontinuierlich bzw. metrisch skaliert und annäherungsweise normalverteilt sein, jedoch können sie auch dichotom als Dummy-Variablen skaliert sein. Zu der multivariaten Normalverteilungsannahme gehört auch,

---

<sup>10</sup> Vgl. dazu Urban/Mayerl 2018: 196-223 und die dortigen Hinweise zum Einsatz der Residuenanalyse für die Überprüfung der Linearitätsannahme sowie Tipps für die Überführung von Modellen mit intrinsischer Linearität in Modelle mit offenen Linearbeziehungen.

<sup>11</sup> Prädiktoren auf der L1-Analyseebene werden "X" genannt. Prädiktoren auf der L2-Ebene sind "Z"-Prädiktoren und Prädiktoren auf der L3-Ebene werden als "W"-Prädiktoren bezeichnet.

dass die dichotomisierten Variablenwerte möglichst balanciert verteilt sind (d.h. bei Werten von 0/1 mit annäherungsweise 50% zu 50% aller Fälle). Sollten diese Verteilungsannahmen nicht gegeben sein, so ist dennoch eine erfolgreiche Schätzung der betreffenden Mehrebenen-Modelle möglich, allerdings sollte dazu unbedingt das robuste Maximum-Likelihood Schätzverfahren (MLR) von Mplus (s. Pkt. 4) eingesetzt werden.

(4) robustes ML-Schätzverfahren (MLR)

Wenn eine oder mehrere der Modellvariablen die Verteilungsvoraussetzungen nicht erfüllen (vgl. Pkt. 3), sollte die Modellschätzung mit einem robusten Maximum-Likelihood-Verfahren durchgeführt werden (in Mplus ist das die Schätzung "MLR", die dort auch als Voreinstellung für MLA benutzt wird).<sup>12</sup> Damit können insbesondere unverzerrte Standardfehler auch bei Abweichungen von den Verteilungsannahmen (s. Pkt. 3) geschätzt werden. Leider können nach Anwendung der MLR-Schätzmethode die damit erzielten Fit-Werte nicht mehr mittels Chi-Quadrat-Differenzentest mit den Fits anderer Modelle verglichen werden. Mehr zur ML(robust)-Schätzmethode findet sich in Urban/Mayerl 2014: 140-145. Wird die MLR-Schätzmethode für Mehrebenen-Analysen (MLA) eingesetzt, so bezeichnen wir dies als MLR/MLA-Verfahren.

(5) ausreichende Fallzahlen

Ausreichende Fallzahlen sind für eine hochwertige MLA-Schätzung unerlässlich. Häufig wird in der Forschung zur Festlegung der benötigten Fallzahlen die 30/30-Regel benutzt (mindestens 30 Makro-Einheiten mit mindestens 30 Mikro-Einheiten pro Makro-Cluster). Aber aufgrund der Ergebnisse neuerer Forschung scheint die 30/30-Regel zu undifferenziert und häufig auch schlichtweg falsch zu sein (vgl. Hox et al. 2018: 212-234). Generell gilt, dass für eine qualitativ hochstehende MLA (nur gering verzerrte Schätzwerte und eine hohe Teststärke) eine große Anzahl von Clustern wichtiger ist als eine große Anzahl von Fällen pro Cluster. Und Simulationsstudien zeigen, dass 50 oder noch mehr L2-Cluster notwendig sind, um die Standardfehler einer MLA-Schätzung akkurat zu ermitteln (vgl. Maas/Hox 2005; Paccagnella 2011). Auch in einer Studie mit logistischen Mehrebenen-Modellen konnte gezeigt werden, dass ein Minimum von 50 Makro-Einheiten und 40 Mikro-Einheiten (pro Makro-Einheit) notwendig ist, um kleine fixierte Effekte ( $OR=1.70$ ) in einem einfachen Modell akkurat zu schätzen. Allerdings zeigte die Studie auch, dass dann, wenn die Modelle komplexer werden und cross-level Effekte akkurat geschätzt werden sollen, ca. 100 Makro-Einheiten und ca. 80 Mikro-Einheiten (pro Makro-Einheit) notwendig werden können (Schoeneberger 2016). Auf jeden Fall scheint eine qualitativ akzeptable Erforschung von Mehrebenen-Modellen zumindest 50 Cluster mit je 20 L1-Fällen zu erfordern, wobei die Empfehlung auch auf 100/10 steigen kann, wenn insbesondere der random Teil der Schätzung (Varianzen, Kovarianzen) erforscht werden soll (Hox et al. 2018: 216).<sup>13</sup> Denn zu geringe Fallzahlen (pro Cluster) reduzieren die statistische Teststärke (=die Wahr-

---

<sup>12</sup> Diese Voreinstellung kann in Mplus durch eine konventionelle ML-Schätzung ersetzt werden mittels der Anweisung: "analysis: estimator=ml;"

<sup>13</sup> Kleinere Fallzahlen scheinen nur dann ausreichend für die Schätzung von MLA-Modellen zu sein, wenn Bayes-Schätztechniken eingesetzt werden (Hox et al. 2018: 215).

scheinlichkeit, einen wahren oder einen signifikanten Effekt zu entdecken) und nicht ausreichende Fallzahlen von L2-Einheiten erhöhen die Typ-1 Fehlerrate bei fixierten L2-Effekten (= das Risiko, einen falschen Effekt zu entdecken) (vgl. Moineddin et al. 2007).<sup>14</sup> Wenn allerdings die Fallzahlen deutlich unter den hier genannten Mindestgrößen liegen, sollte anstelle einer MLA eine MLA/OLS-Regressionsanalyse mit L2-Dummies (als Haupt- und Interaktionseffekte) durchgeführt werden (vgl. dazu Kap. A3, Pkt. 7).

(6) fehlende Werte sind “missing at random” (MAR)

In Mplus wird angenommen, dass im Datensatz zur Durchführung einer MLA mittels MLR (=robuste Maximum Likelihood-Schätzung, s.o.) die fehlenden Werte bzw. “Missings” durch den Ausfallmechanismus “missing at random” (MAR) zustande kommen (als Voreinstellung). Für andere Ausfallmechanismen (“missing completely at random” oder “not missing at random”) bietet Mplus spezielle Strategien an. Zur Erläuterung der verschiedenen Ausfallmechanismen bei der Entstehung von Missings vgl. Urban/Mayerl 2018: 445-447.

(7) Abwesenheit von Multikollinearität

Im Idealfall sollte in einer MLA jede X/Z/W-Variable keine andere unabhängige Variable auf der gleichen Analyseebene entscheidend beeinflussen. Es sollten also keine gravierenden Kovarianzen zwischen den X-Variablen auf L1, zwischen den Z-Variablen auf L2 und/oder zwischen den W-Variablen auf L3 bestehen. Hohe Kovarianzen zwischen den unabh. Variablen können verzerrte Parameterschätzwerte (insbesondere Verzerrungen bei Schätzung von Varianzen und Standardfehlern) zur Folge haben. Leider kann auch eine Zentrierung der unabh. Variablen nicht in jedem Fall, wie häufig angenommen wird, eine vorhandene Multikollinearität beseitigen (vgl. Urban/ Mayerl 2018: 248-252). Denn in der Analyse ist zwischen einer essenziellen Multikollinearität (d.h. einer Kovarianz zwischen Variablen mit selbstständigen Messbereichen) und einer nicht-essenziellen Multikollinearität (z.B. aufgrund einer Kovarianz zwischen Variablen und daraus abgeleiteten Interaktionsvariablen) zu unterscheiden. Nur die essenzielle Multikollinearität stellt für die Modellschätzung ein Problem dar und sollte in der MLA beseitigt oder abgemildert werden. Für weitere Informationen zur Entstehung, Diagnose und Beseitigung von Multikollinearität vgl. Urban/Mayerl 2018: 231-252.

(8) Abwesenheit von Ausreißer-Werten (outliers)

Insbesondere bei kleinen Fallzahlen (auf welcher Ebene auch immer) können einzelne extreme Ausreißer-Werte die ML-geschätzten Modellparameter stark verzerren. Deshalb sollten die benutzten Datensätze schon zu Anfang einer MLA hinsichtlich einzelner Ausreißer-Werte kontrolliert werden und müssten diese (wenn möglich) aus dem Datensatz entfernt werden. In welcher Weise dies geschehen kann, wird in Urban 2018: 177-195 erläutert.

---

<sup>14</sup> Generell betrachtet ist es empfehlenswert, die für eine MLA erforderlichen Fallzahlen über Simulationsstudien und Teststärke-Untersuchungen abzusichern (Hox et al. 2018: 212f, 219-221, 229f, 233f; Muthen/Muthen 2002).

(9) modellübergreifende Unabhängigkeit der Beobachtungen

Auch jenseits der im Mehrebenen-Modell berücksichtigten Kontextabhängigkeit der Beobachtungen sollten die Beobachtungswerte unabhängig voneinander sein und sollten nicht durch zusätzliche, im Modell unberücksichtigt bleibende Kontexte miteinander verbunden sein. Wenn z.B. die Lebenszufriedenheit von Schüler:innen untersucht werden soll und wenn dabei die Zugehörigkeit zu verschiedenen Schulklassen als L2-Ebene in einem Mehrebenen-Modell berücksichtigt wird, sollte nicht durch die unberücksichtigt gebliebene Zugehörigkeit der Schüler:innen zu verschiedenen politischen Jugendvereinen eine Kontextabhängigkeit der Lebenszufriedenheitswerte entstehen, die in der statistischen Analyse die Unabhängigkeit der Beobachtungen aufheben würde.

(10) keine ebenenspezifischen Residuen-Kovarianzen

Wenn ein Modell mindestens zwei Random-Komponenten enthält (z.B. ein Random-Intercept und einen Random-Slope), und wenn diese Komponenten auf der gleichen Ebene liegen (z.B. auf L2), dann könnten die Residuen dieser Random-Komponenten miteinander kovariieren. So zeigt z.B. die Abbildung 6.1 (zu Modell 06) die Residuen  $u_{0j}$  und  $u_{1j}$ , die in Form einer Kovarianz miteinander verbunden sind.

In Mplus werden jedoch in der Voreinstellung keine Kovarianzen geschätzt und es wird angenommen, dass diese Kovarianzen einen Wert von 0 haben. Deshalb müssen in Mplus die oben angesprochenen Kovarianzen separat angefordert werden, um sie in die Modellschätzung aufzunehmen. Bei der Interpretation dieser Kovarianzen ist darauf zu achten, dass es sich bei der Kovarianz zwischen den Residuen  $u_{0j}$  und  $u_{1j}$  um die Kovarianz zwischen der Varianz der abh. Variablen (Random Intercept) und der Varianz eines Steigungsfaktors (Random-Slope) handeln kann. Dies ist z.B. bei dem zuvor genannten Modell 06 der Fall. Weitere Informationen zur Residuen-Kovarianz finden sich in diesem Skript in Kap. B3.

(11) keine Random-Effekte auf der höchsten Analyseebene

In Mehrebenen-Modellen können auf der höchsten Ebene keine Random-Slopes bzw. Random-Effekte spezifiziert werden. Das gilt auch für cross-level Effekte. Wenn ein Modell z.B. nur zwei Ebenen umfasst, müssen die L2-Kontextvariablen "Z" dementsprechend feste/fixed Effekte und keine random Effekte aufweisen.



### A3 Welche Vorteile hat eine Mehrebenen-Analyse?

Im Folgenden wollen wir stichpunktartig die Vorteile einer statistischen Mehrebenen-Analyse (MLA=Multi-Level-Analysis) im Unterschied zur Analyse von Daten ohne Mehr-Ebenen-Differenzierung aufführen.

Eine statistische Mehrebenen-Analyse ist immer dann vorteilhaft, wenn die zu analysierenden Daten hierarchisch geschachtelt (nested) sind, d.h. wenn es z.B. Daten zu individuellen Einstellungen von befragten Personen gibt (wie z.B. zu deren "Arbeitszufriedenheit") und wenn es gleichzeitig Daten zu unterschiedlichen einstellungsrelevanten Kontexten gibt, in denen sich die befragten Personen aufhalten (wie z.B. in Wirtschaftsunternehmen mit unterschiedlichen Strukturen von "Verantwortung am Arbeitsplatz"). Eine MLA sollte also dann ausgeführt werden, wenn es Daten über befragte Personen gibt (Individualebene) und wenn es gleichzeitig auch Makro-Daten gibt, die für alle Mitglieder der entsprechenden Kontexte/ Makroeinheiten gültig sind und die auch die Ausprägungen der Individualdaten beeinflussen können (wie z.B.: "für alle Beschäftigten von Unternehmen A gilt, dass der Grad individueller Verantwortung am Arbeitsplatz überdurchschnittlich hoch ausgeprägt ist"). Wenn solche Datenstrukturen vorliegen, dann sollte die statistische Analyse dieser Daten in Form einer Mehrebenen-Analyse durchgeführt werden. Nur wenn es keine Varianz zwischen den Makroeinheiten gäbe, wäre eine spezielle MLA überflüssig und sollte eine MLA/OLS-Schätzung durchgeführt werden (vgl. dazu auch Kap. A1).

Wir wollen das hier Gemeinde anhand der folgenden Abbildung A3.1 verdeutlichen (nach: Eid et al. 2013: 729).

Die Abbildung zeigt in ihrem oberen Teil den linearen Zusammenhang zwischen Arbeitszufriedenheit (abh. Y-Variable) und dem Grad individueller Verantwortung am Arbeitsplatz (unabh. X-Variable) ohne eine Differenzierung nach Wirtschaftsunternehmen. Der lineare Zusammenhang zwischen beiden Variablen ist stark negativ ausgeprägt und beträgt: -0,69 (stand. Regressionskoeffizient).

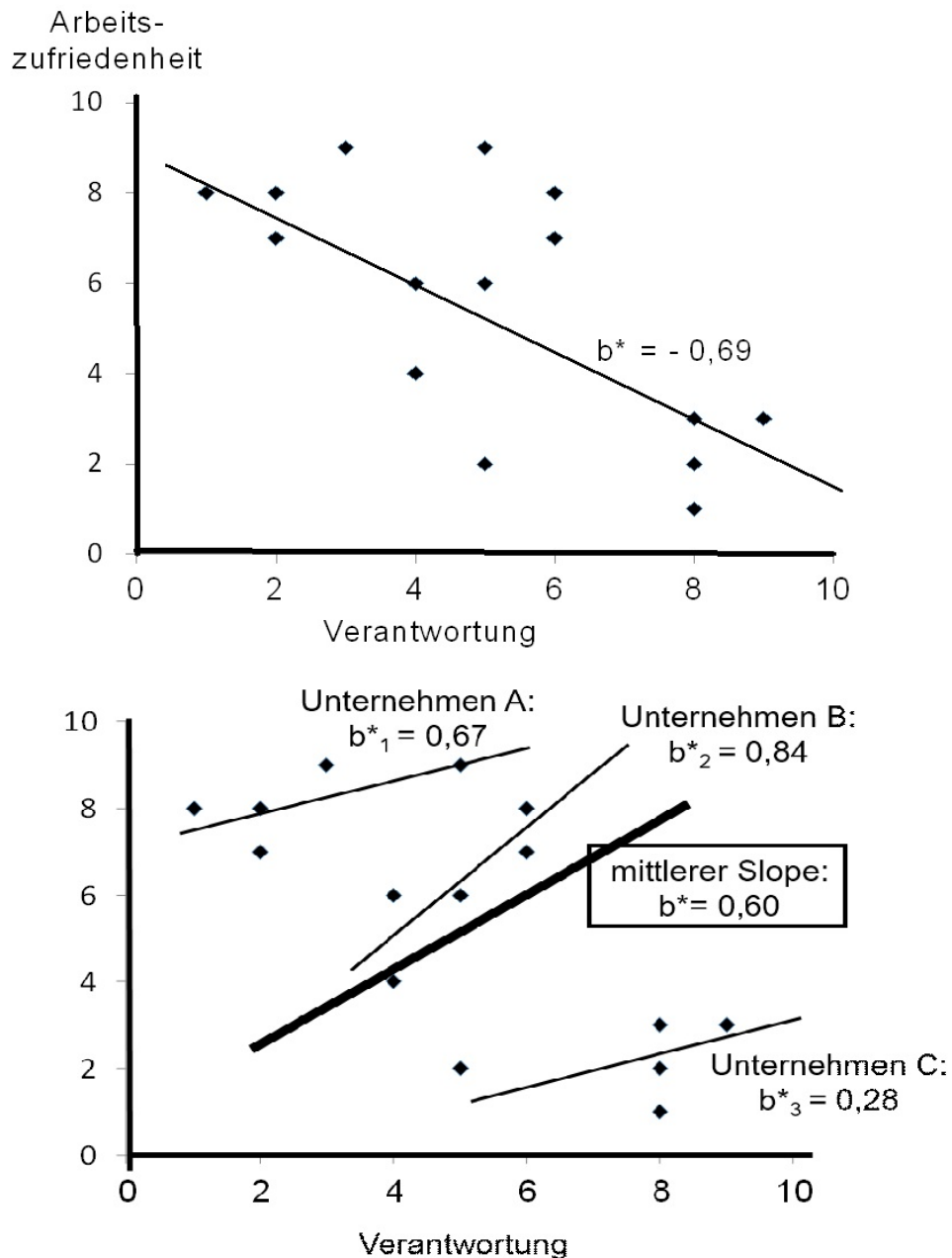
Im unteren Teil der Abbildung werden die gleichen Daten nach den drei Unternehmen, in denen die Daten gesammelt wurden, aufgeschlüsselt. Es wird angenommen, dass der Y/X-Zusammenhang je nach beschäftigendem Unternehmen unterschiedlich ausfällt, d.h. dass die Intercepts je nach Unternehmen variieren ("random intercepts") und dass auch die Steigungskoeffizienten je nach Unternehmen variieren ("random slopes"). Wie die Abbildung sehr deutlich zeigt, ist mit diesen Annahmen der Zusammenhang zwischen Y und X nicht mehr negativ, sondern in jedem einzelnen Unternehmen positiv, und auch der mittlere lineare Zusammenhang ist dann ebenfalls positiv und beträgt +0,60 (stand. Repr.Koeffizient).

Mit einer MLA und den genannten Annahmen können also die unterschiedlichen Gruppenniveaus hinsichtlich der in jedem Unternehmen zu messenden Arbeitszufriedenheit berücksichtigt werden, und können auch gleichzeitig die je nach Unternehmen variierenden Slopes beim Zusammenhang zwischen Arbeitszufriedenheit und Verantwortung differenziell kalkuliert werden. Dadurch ergibt sich eine mittlere Regressionslinie, die anders als bei der MLA/OLS-Schätzung positiv ansteigt und nicht mehr negativ abfällt und deren absoluter Wert ( $b_{YX}^* = r_{YX}$ ) auch deutlich niedriger ist (+0,60 anstelle von -0,67).

Eine MLA/OLS-Schätzung, die ohne Berücksichtigung der Ebenendifferenzen verfährt und deshalb bei ihrer Schätzung alle Datenpunkte in gleicher Weise zur Festlegung der Regressionsgeraden benutzt, würde also den Y/X-Zusammenhang bzw. den Zusammenhang zwischen Arbeitszufriedenheit und Verantwortung eindeutig falsch bzw. stark verzerrt und mit unzutreffender Richtungsangabe schätzen.

Methodologisch betrachtet handelt es sich bei der Nicht-Berücksichtigung von Kontextvariablen um einen Spezifikationsfehler. Und wenn die Kontextvariablen weggelassen werden, obwohl diese eine wichtige Bedeutung für den spezifizierten Modellzusammenhang haben, können Individualeffektschätzungen schnell inflationiert sein.

Abb. A3.1: Zwei-Ebenen-Struktur in einem Datensatz (fiktiv) mit linearer Abhängigkeit (nach: Eid et al. 2013: 729)



Im einzelnen bietet eine Mehrebenen-Analyse von Daten, die (wie oben beschrieben) hierarchisch geschachtelt sind, folgende Vorteile:

1. Fast alle klassischen statistischen Analysen verlangen die Unabhängigkeit der gemessenen Daten zwischen den Messeinheiten. Die Daten von Person A sollten also nicht durch die Daten von Person B bestimmt sein. Wenn aber z.B. die Personen A, B, C usw. dem gleichen Sportverein angehören, bestimmt dessen Angebot auch die Sportaktivitäten aller Vereinsmitglieder, d.h. die Vereinsmitglieder sind hinsichtlich ihrer sportlichen Aktivitäten nicht unabhängig voneinander. Eine MLA kann solche ebenen-übergreifenden Abhängigkeiten in ihren statistischen Operationen explizit berücksichtigen.
2. Daten, die nicht unabhängig voneinander sind, erzeugen verzerrte Standardfehler, deren Ausmaß oftmals zu gering ausfällt (negativ biased). Mit kleineren Standardfehlern wird in statistischen Tests das Risiko für inflationierte Fehler der 1. Art erhöht, sodass Signifikanztests häufig irrtümlicherweise signifikante Ergebnisse vermelden. Es erfolgt mithin eine Überschätzung der Signifikanzen bzw. eine alpha-Fehler-Inflation (viel zu viele Parameter werden viel zu hoch signifikant). Auch die diesbezüglichen Konfidenzintervalle werden oftmals als zu gering bzw. als zu eng geschätzt, sodass die geschätzten Parameterwerte schnell und fälschlich im 5%- oder 1%-Bereich der Signifikanzkurve zu liegen kommen. Diese verfälschten Testergebnisse können durch eine MLA vermieden werden.
3. Verzerrte Standardfehler (vgl. Pkt. 2) signalisieren, dass die Zuverlässigkeit von Modellschätzungen gering sein kann und deshalb Modellparameter falsch geschätzt sein können. Dazu kann auch gehören, dass bestimmte Modellparameter, wie z.B. spezielle Regressionsparameter, fälschlicherweise in die Modellschätzung aufgenommen oder fälschlicherweise ausgelassen wurden.
4. In Modellen, mittels denen das Ausmaß der Abhängigkeit bestimmter Y-Variablen (z.B. von "individueller Lebenszufriedenheit") von bestimmten unabh. Variablen auf verschiedenen Ebenen geschätzt werden soll, können die Effekte sowohl separat für Einflüsse auf der Mikro- und der Makroebene als auch für ebenenüberschreitende Effekte (Cross-Level-Effekte) ermittelt werden.
5. Mit einer MLA können vorteilhaft Längsschnittanalysen mit Paneldaten durchgeführt werden: auf der untersten Ebene (L1) sind dann die zeitvarianten Beobachtungswerte angesiedelt (wie z.B. Einstellungsmesswerte zu den Zeitpunkten  $t_0, t_1, t_2, t_3$ ). Die dazugehörigen Personen mit ihren personalen, zeitinvarianten Werten (wie z.B. "Alter" oder "Geschlecht") liegen dann in der MLA auf der übergeordneten Analyseebene (L2). Auf diese Weise können auch Personen, die Missings aufweisen, für deren Eigenschaften (z.B. Einstellungen) also nicht zu jedem Beobachtungspunkt ein gültiger L1-Messwert vorliegt, im Datensatz für eine Panelanalyse mit Mehrebenen-Design beibehalten werden.

Im vorliegenden Skript werden MLA zur Längsschnittanalysen mit Paneldaten nicht behandelt (vgl. dazu Finch/Bolin 2017: 87-96).

6. MLA mit MLR-Schätzverfahren können viele Nachteile von MLAs mit OLS-Schätzung (MLA/OLS-Schätzung) vermeiden. Zu den MLA/OLS-Schätzungen, deren Nachteile ausgeräumt werden können, gehören:
- a) OLS-Schätzungen mit Dummies  
In dieser OLS-Modellierung müssen für jede Makroeinheit (d.h. für jede Einheit auf L2- oder höherer Ebene) eine Dummy-Variable und gleichzeitig auch Interaktionsvariablen spezifiziert werden (denn jeder Effekt muss auch als Interaktionseffekt geschätzt werden). Deshalb kann diese Form der MLA nur benutzt werden, wenn wenige Makroeinheiten und wenige Prädiktoren zu untersuchen sind. Ansonsten explodiert die Anzahl der zu schätzenden Parameter schnell ins Unüberschaubare. MLA mit MLR-Schätzung vermeiden diese Inflation an zu schätzenden Parametern. Sie können auch bei Analysen mit vielen Makro-Einheiten eingesetzt werden.
  - b) OLS-Schätzungen mit disaggregierten Makromerkmalen (nach Hadler 2004)  
In dieser OLS-Modellierung werden die Makromerkmale als Individualmerkmale behandelt und wird somit jedes Makromerkmal vielen Individualeinheiten zugeordnet. Dadurch entstehen zwei gravierende Nachteile dieses Verfahrens. Zum einen werden die Standardfehler stark erniedrigt und zum anderen werden die OLS-Schätzer nur verzerrt ermittelt (vgl. auch Abb. A3.1 in Kap. A3).<sup>15</sup> Zu geringe Standardfehler ergeben jedoch häufig falsche Signifikanzen und verzerrte Parameterschätzungen verfälschen schnell die substantielle Bedeutung von Modellanalysen. MLA mit MLR-Schätzern können diese Fehler vermeiden.
  - c) OLS-Schätzungen mit separierten Modellen  
In dieser OLS-Modellierung wird für jede Makroeinheit ein separat-komplettes Modell geschätzt. Zur Schätzung einer Vielzahl von Modellen müssen aber auch die Fallzahlen dementsprechend groß sein. Zudem müssen in dieser Forschungsstrategie die Differenzen zwischen den einzelnen Modellschätzungen in qualitativer Weise als Kontexteffekte interpretiert werden, da in den Schätzungen keine Kontextdifferenzen berücksichtigt werden. Welche Differenzen aber als substantiell bedeutsam interpretiert werden, liegt im Ermessensspielraum der Forschenden. Eine MLA/MLR-Schätzung liefert demgegenüber Schätzwerte für die Kontexteffekte und Informationen über deren Relevanz (Signifikanz) in einer einzigen Modellschätzung.
7. Mehrebenen-Analysen mit MLR-Schätzverfahren (wie mit Mplus möglich) können auch benutzt werden, um eine große Vielzahl von Modelltypen mit Mikro/Makro-Differenzen im Kontext von generalisierten linearen Modellen (GLM) zu untersuchen. Zu den diesbezüglichen Modellen gehören: logistische Regressionsmodelle (bi- und multinomial), Poisson Regressionsmodelle, Modelle mit Zähldaten, Survival-Modelle, Modelle mit latenten Variablen, Strukturgleichungsmodelle, Mixture-Modelle, Latent-Class-Modelle, Wachstumskurven-Modelle u.a. Die aufgeführten Modelle werden im vorliegenden Skript nicht erörtert (vgl. dazu Finch/Bolin 2017, Heck/Thomas 2015, Hox et al. 2018).

---

<sup>15</sup> In einer Beispielsrechnung ermittelte Hadler (2004) mit einer MLA/OLS-Schätzung einen deutlich zu niedrigen Standardfehler von 0,036 (anstatt eines Wertes von 0,387), und eine verzerrte Parameter-Schätzung von 4,45 (anstatt eines Wertes von 2,74), wenn fälschlicherweise ein MLA/OLS-Schätzer eingesetzt wurde.

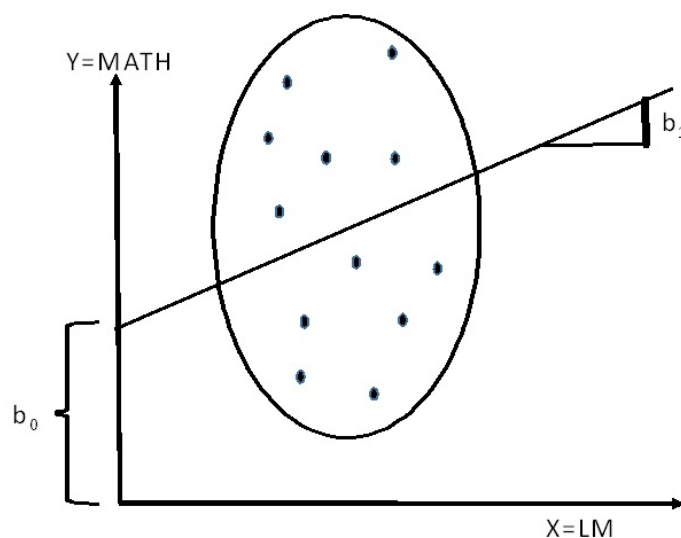
## B Einige ausgewählte Grundlagen der Mehrebenen-Analyse

### B1 Das Analysemodell

Wir wollen im Folgenden von einem leicht zu verstehenden Beispiel ausgehen: Nehmen wir dazu an, dass wir Daten über 12 Schüler<sup>16</sup> aus 4 unterschiedlichen Schulklassen in 2 verschiedenen Schulen besitzen. Wir haben also Daten, die auf drei verschiedenen Ebenen bzw. Level (L) liegen: L1: Schüler, L2: Schulklassen, L3: Schulen. Für eine reale Mehrebenen-Analyse werden natürlich mehr Fälle auf jeder Ebene der Analyse benötigt (vgl. dazu Kap. A2, Pkt. 5). Für unser Beispiel reichen jedoch die Daten über 12 Schüler und deren Schulklassen sowie deren Schulen aus. Zudem gehen wir vereinfachend davon aus, dass sich alle im Folgenden gezeigten Abbildungen auf Modelle mit nur zwei Variablen beziehen, sodass es sich bei den hier erläuterten Mehrebenen-Analysen ausschließlich um bivariate Modelle handelt.

In Abbildung 1a wird das bekannte Modell einer bivariaten Regressionsschätzung veranschaulicht. Die zwölf Schüler haben jeweils eine spezifische Kompetenz zur Lösung altersgerechter mathematischer Übungsaufgaben ( $Y=MATH$ ) und jeweils eine unterschiedlich hoch ausgeprägte Lernmotivation ( $X=LM$ ). Im Koordinatensystem von Abb. 1a symbolisiert jeder Punkt die Lage eines bestimmten Schülers. Diese Lage wird durch seinen MATH-Wert und seinen LM-Wert bestimmt. Der lineare Zusammenhang zwischen MATH und LM über alle Schüler hinweg wird durch den Verlauf einer Regressionsgeraden veranschaulicht. Deren Schnittpunkt mit der Y-Linie liegt beim Intercept " $b_0$ ", dessen Wert die für alle Schüler zu erwartende Mathematikleistung bezeichnet. Dieser Wert ist dann zu erwarten, wenn der Effekt von LM unberücksichtigt bleibt oder gleich null ist. Zudem weist die Regressionsgerade in Abb. 1a eine Steigung von  $b_1$  auf. Je größer diese Steigung ist, desto stärker ist der zu erwartende Einfluss von LM auf MATH.

Abb. 1a: OLS-Regression

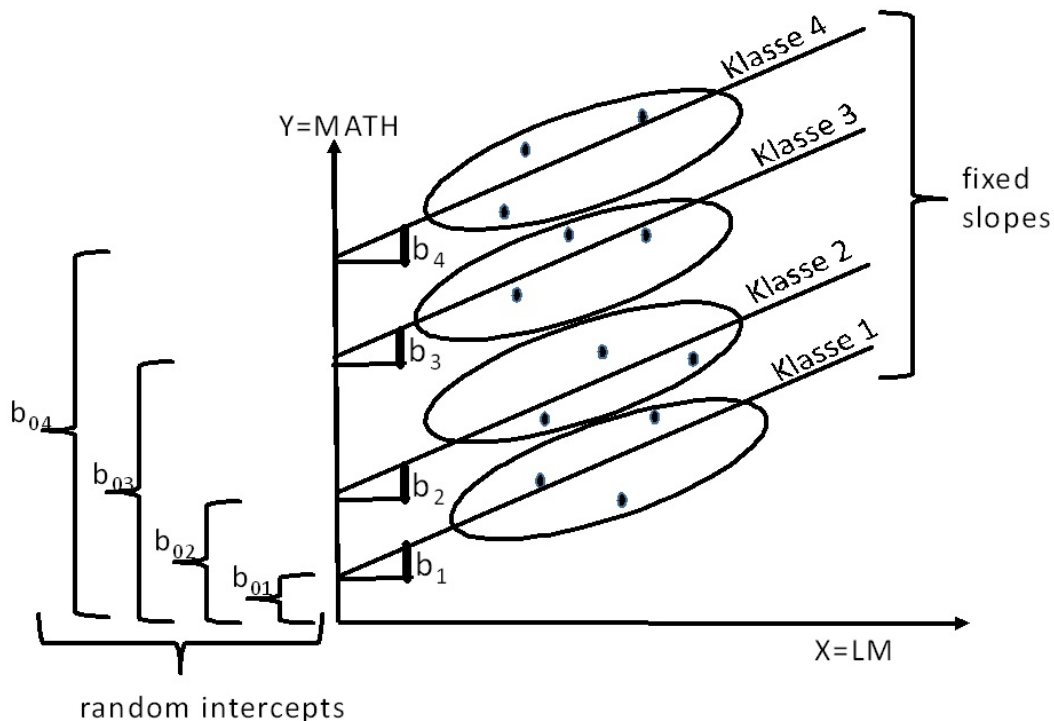


<sup>16</sup> Das hier und im Folgenden benutzte generische Maskulinum beschreibt die Verwendung der männlichen grammatikalischen Form unabhängig vom Geschlecht der beschriebenen Person.

Die Abbildung 1a betrifft eine Analyse auf der Mikro-Ebene eines Mehrebenen-Modells. Diese Ebene bezieht sich oftmals auf die Relationen und Zustände der zu untersuchenden Akteure und wird hier als "Level 1" oder "L1" bezeichnet. Die Ebene L1 ist die unterste, nicht weiter aufzulösende Analyseebene in einem hierarchisch aufgebauten, sozialwissenschaftlichen Mehrebenen-Modell.

In Abbildung 1b wird der Übergang von einer L1-Analyse zu einer L2-Analyse dargestellt. Alle 12 Schüler mit ihren Werten für MATH und LM befinden sich im Koordinatensystem noch an den gleichen Positionen wie in Abb. 1a. Jedoch gehören jeweils drei Schüler zu einem Klassenverband und liegen deshalb in der Abbildung innerhalb eines von vier Ovalen, die jeweils den Klassenverband einer Schulklasse symbolisieren. Somit kann die Regression von MATH auf LM für die Schüler einer jeden Klasse separat durchgeführt werden und es werden vier separate Regressionsgeraden geschätzt, die aufgrund einer vereinfachenden Annahme (die vom Analytischen zu treffen ist) alle die gleiche Steigung aufweisen ("fixed slopes"). Die Steigungen aller vier Geraden ( $b_1$  bis  $b_4$ ) sind gleich groß. Allerdings besitzen die vier Regressionsgeraden vier unterschiedliche  $b_0$ -Intercepts ("random intercepts"), sodass deutlich wird, dass die Mathematikkompetenz nicht nur von der Leistungsmotivation der Schüler, sondern auch von Eigenschaften des jeweiligen Klassenverbands abhängt.

Abb. 1b: L1-Regressionen nach Klasseneinteilung

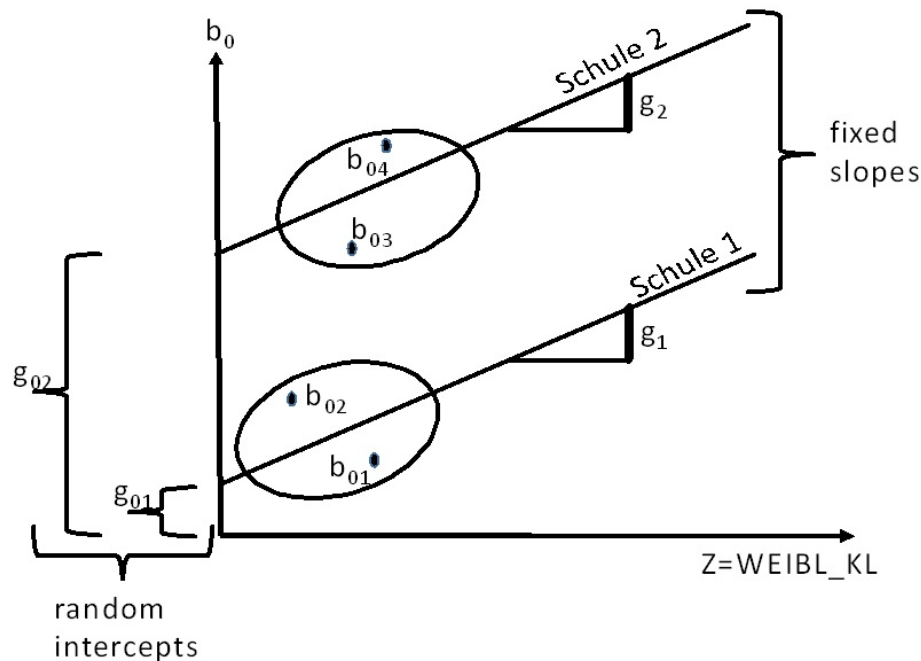


Die Clustereinteilung nach Schulklassen (deutlich in Abb. 1b zu erkennen) führt also dazu, dass das Intercept  $b_{0j}$  ein gewisses Ausmaß an Variation aufweist und es gilt: je größer  $b_{0j}$ , umso besser (oder im Koordinatensystem: höher/größer) ist die Mathematik-Leistung der entsprechenden Schulklasse (vgl. Abb. 1b). Will man also herausfinden, wie stark eine L2-Variable (wie z.B. der Anteil an Schülerinnen in jeder Klasse: WEIBL\_KL) zum Ausmaß von MATH beiträgt, kann man in einer Regressionschätzung die vier Intercepts  $b_{0j}$  als Werte einer abhängigen Variablen

betrachten, die von einem L2-Prädiktor beeinflusst werden. In unserem stark vereinfachten Beispiel wären mithin die vier  $b_{0j}$ -Werte als Ausprägungen der abh. L2-Variablen “Mathematikleistung pro Klasse” zu verstehen. Und als unabhängige L2-Prädiktorvariable könnte die Variable “Anteil an Schülerinnen in jeder Klasse” oder “WEIBL\_KL” benutzt werden.

Das Ergebnis einer Regressionsschätzung mit den oben genannten Variablen zeigt Abb. 1c. Dort werden die klassenspezifischen Intercepts  $b_{0j}$  als Werte einer abh. Variablen verstanden und als L2-Prädiktor die Variable “Anteil weiblicher Schülerinnen in der Schulklasse” (WEIBL-KL) benutzt.

Abb. 1c: L2-Regressionen nach Schuleinteilung



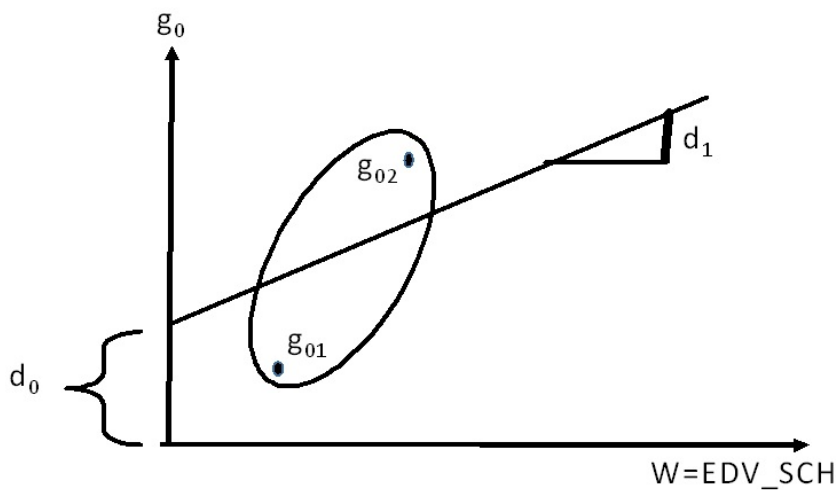
Wie Abbildung 1c zeigt, befinden sich nunmehr im regressionsrelevanten Koordinatensystem nur noch vier Datenpunkte. Dies deshalb, weil jetzt die Analyse auf L2-Niveau durchgeführt wird. Als abh. Y-Werte stehen nur noch die Werte von vier Intercepts zur Verfügung (für jede Schulklasse ein  $b_{0j}$ -Intercept). Und weil es in der Untersuchung zwei Schulen gibt, können die  $b_{0j}$ -Intercepts, wie in Abb. 1c gezeigt, auch nach Schulen geclustert werden..

Mit den  $b_{0j}$ -Intercept-Werten der Schulklassen kann für jede der zwei Schulen ein Regressionsmodell geschätzt werden, bei dem die L2-Werte von  $b_{0j}$  auf die Effekte des L2-Prädiktors zurückgeführt werden. In unserem Beispiel ist dieser Prädiktor (auch Z-Variable genannt) die Anzahl von Schülerinnen pro Klasse (WEIBL\_KL, s.o.). Mit den beiden Regressionsschätzungen (pro Schule eine Schätzung) von  $b_{0j}$  auf WEIBL\_KL erhalten wir wieder zwei Regressionsgeraden mit je einem Intercept ( $g_{0j}$ ) (vgl. dazu Abb. 1c). Die zwei Intercepts dieser beiden Schätzungen ( $g_{01}$  und  $g_{02}$ ) benennen die Mathematikleistung pro Schule ohne Berücksichtigung der WEIBL\_KL-Variable.

In Abbildung 1d gibt es im Oval nur noch die beiden schulspezifischen  $g_{0k}$ -Intercepts, die zuvor (vgl. Abb. 1c) regressionsanalytisch geschätzt wurden. Analytisch betrachtet, liegen sie auf der L3-Ebene. Ihre Abhängigkeit von einem Prädiktor kann somit nur ermittelt werden, wenn

auch die unabh. Prädiktorvariable auf der L3-Ebene angesiedelt ist. In unserem Beispiel sei das die Ausstattung einer Schule mit modernen EDV-Geräten zur Unterstützung der Mathematiklehre “EDV\_SCH” (auch W-Variable genannt). Wie in Abb. 1d zu erkennen ist, gibt es aufgrund unseres sehr kleinen Samples von nur 12 Schülern aus 2 Schulen für die Regressions-schätzung nur zwei Intercept-Werte (pro Schule einen Intercept-Wert). Die Schätzung mit diesen beiden g-Intercepts als abh. Variable und mit EDV\_SCH als unabh. Prädiktorvariable erbringt den Intercept-Wert  $d_0$  und den Steigungskoeffizient von  $d_1$ . Dieser beziffert den Effekt von EDV\_SCH auf  $g_0$  und damit indirekt auch auf  $b_0$ . Die Schätzung bewegt sich aber allein auf Analyseebene L3.

Abb. 1d: L3-Regression

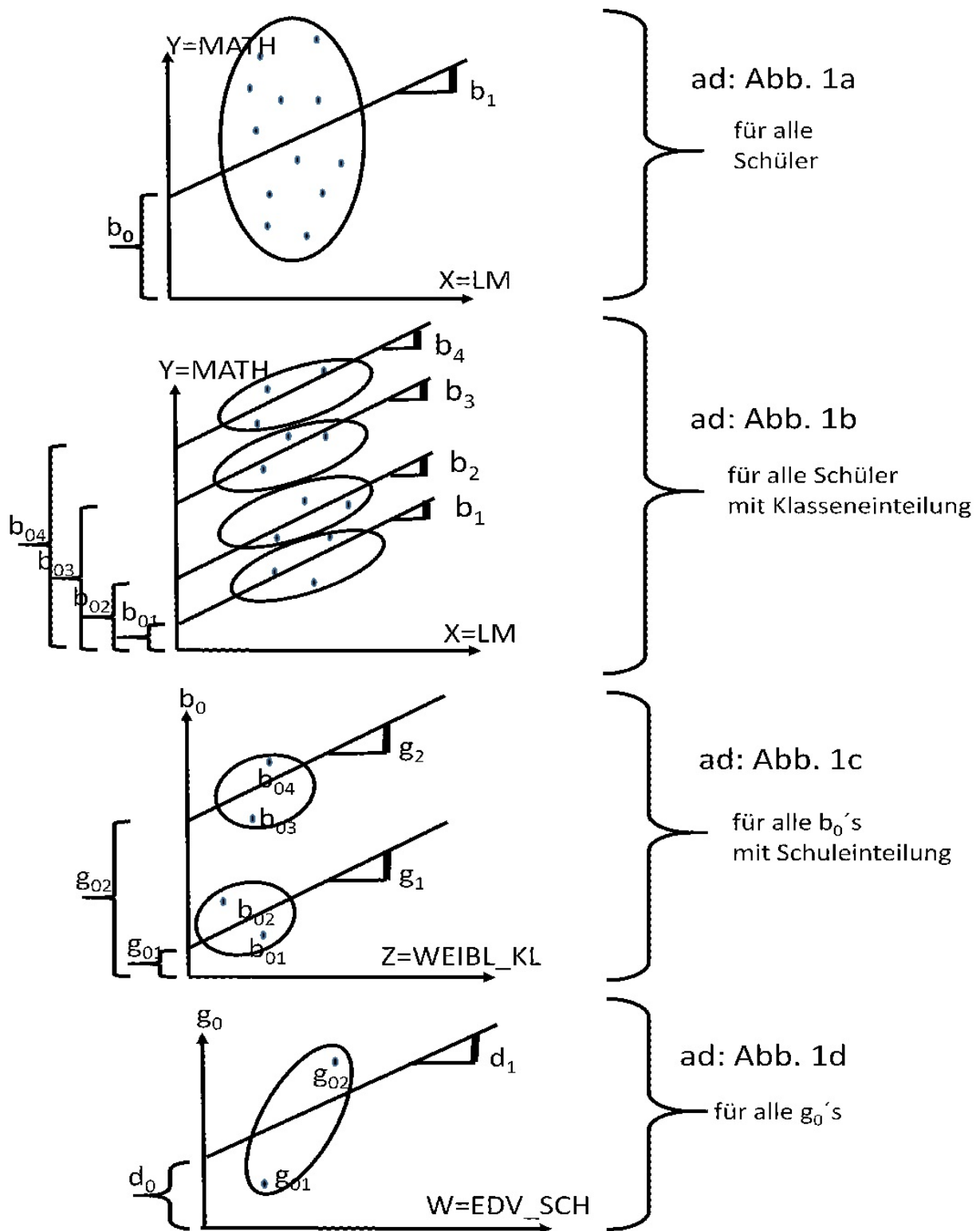


Es lässt sich jetzt also leicht erkennen: Der eigentliche Trick bei der Mehrebenen-Analyse besteht darin, auf jeder Ebene der Analyse spezifische Daten für die abhängige Variable (bei uns: MATH) via Regressionsanalyse schätzen zu können. Diese Daten müssen von Ebene L1 bis Ebene L3 auseinander ableitbar sind. Bei uns waren das die vier b-Intercepts, die in zwei g-Intercepts und dann in ein d-Intercept überführt werden konnten.

Dies lässt sich auch noch einmal besonders gut anhand der folgenden Abbildung 1e erkennen, in der die verschiedenen Stufen der Analyse in einer Grafik gegenübergestellt werden. Zudem verweist die Abbildung auch noch einmal auf die verschiedenen ebenenspezifischen Prädiktoren (LM, WEIBL\_KI, EDV\_SCH), die auf jeder Ebene als Beschriftung der X-Achse gut erkennbar sind.



Abb. 1e: Zusammenschau von Abb. 1a bis Abb. 1d



Die drei Ebenen der Modellschätzung können nicht nur graphisch sondern auch rein algebraisch (ohne Zeichnungen) anhand von drei Lineargleichungen deutlich gemacht werden:

$$\text{L1-Gleichung: } Y_{ijk} = b_{0jk} + b_{1jk}X_{ijk} + e_{ijk}$$

$$\text{L2-Gleichung: } b_{0jk} = g_{00k} + g_{01k}Z_{jk} + u_{0jk}$$

$$\text{L3-Gleichung: } g_{00k} = d_{000} + d_{001}W_k + v_{00k}$$

Die L1-Gleichung benennt den Sachverhalt, den wir schon anhand der Abbildungen 1a und 1b erläutert haben: Es wird nach dieser Gleichung zur Prognose von Y für jeden Schüler (Subskript: i) aus jeder Schulklasse (Subskript: j) von jeder Schule (Subskript: k) eine Lineargleichung mit einem Intercept ( $b_0$ ) und einem Slope ( $b_1$ ) für den Effekt von X geschätzt.

Die L2-Gleichung zeigt sodann, dass die vier klassenspezifischen b-Intercepts zu abhängigen Variablen werden können und für jede der beiden Schulen eine Schätzgleichung aufgestellt werden kann, mit der jeweils ein abhängiges  $b_0$ -Intercept mit dem neu zu schätzenden  $g_0$ -Intercept sowie mit dem neu zu schätzendem  $g_1$ -Slope und mit der empirisch beobachteten Z-Variablen zu prognostizieren ist (vgl. Abb. 1c).

Die L3-Gleichung berücksichtigt als Ausprägungen einer abh. g-Variablen die beiden  $g_0$ 's aus Abb. 1c und ermittelt deren Werte mittels einer Linearschätzung mit den Parametern  $d_0$  und  $d_1$  sowie den beobachteten Werten der W-Variablen "EDV\_SCH" (vgl. Abb. 1d).

Die drei oben angeführten L1/L2/L3-Gleichungen stehen auch im Zentrum des folgenden Gleichungssystems, das unser Mehrebenen-Übungsmodell komplett beschreibt (die zu schätzenden Parameter werden dabei mit griechischen Buchstaben bezeichnet):<sup>17</sup>

$$\text{L1-Gleichung } Y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{ijk} + e_{ijk} \quad (1f)$$

$$\text{L2-Gleichungen: } \beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z_{jk} + u_{0jk} \quad (1g)$$

$$\beta_{1jk} = \gamma_{10k} \quad (1h)$$

$$\text{L3-Gleichungen: } \gamma_{00k} = \delta_{000} + \delta_{001}W_k + v_{00k} \quad (1i)$$

$$\gamma_{01k} = \delta_{010} \quad (1j)$$

$$\gamma_{10k} = \delta_{100} \quad (1k)$$

$$\text{"fixed part": } \delta_{000}, \delta_{001}, \delta_{010}, \delta_{100}$$

$$\text{"random part": } \text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(v_{00k})$$

In dieser rein algebraischen Beschreibung unseres Mehrebenen-Übungsmodells sind drei neue Gleichungen hinzu gekommen, die wir bislang noch nicht kennengelernt haben. Das sind die Gleichungen 1h, 1j und 1k. Mit ihnen wird beschrieben, dass die zu schätzenden Parameter  $\beta_{1jk}$ ,  $\gamma_{01k}$  und  $\gamma_{10k}$  als feste (fixed) Parameter zu verstehen sind, deren Werte nicht zwischen den

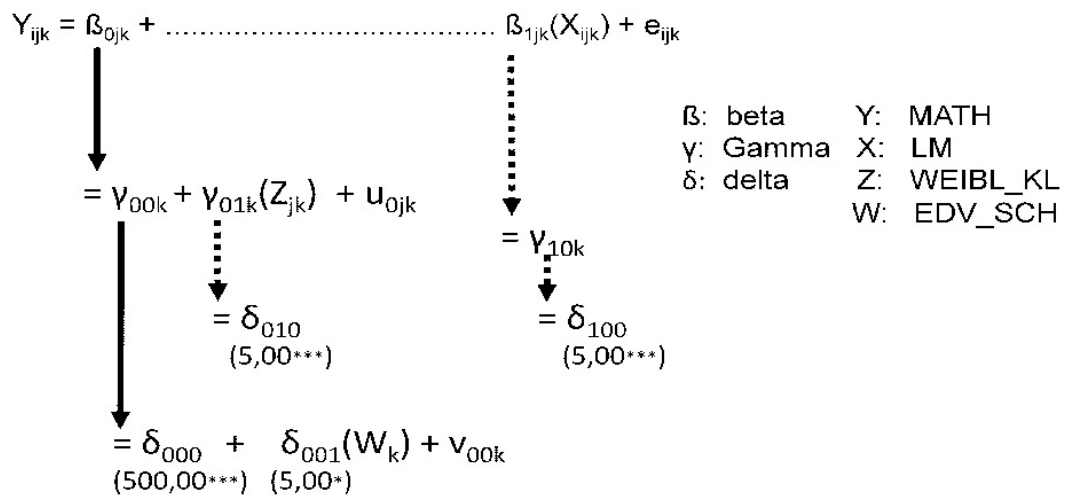
<sup>17</sup> Dieses Gleichungssystem dient Übungszwecken und ist fiktiv. Das gilt auch insbesondere für die Parameterwerte dieses Modells (in: Abb. 1f), die ebenfalls rein fiktiv sind. In seiner Logik entspricht dieses Modell am ehesten dem empirischen Modell 21 (in Kap. D4), in dem allerdings auf jeder Analyseebene je ein zusätzlicher Y/Z/W-Effekt aufgenommen wird.

einzelnen Schulklassen bzw. zwischen den variieren Schulen variieren (mehr dazu im Folgenden).

Das oben abgedruckte Gleichungssystem ist nicht einfach zu verstehen. Wir haben es deshalb noch einmal graphisch aufbereitet und zeigen die Gleichungen in der folgenden Abbildung 1f. Die darin enthaltenen Pfeile symbolisieren keine Kausalitäten, sondern verknüpfen immer zwei Gleichungsseiten miteinander, sodass z.B. die folgende Abbildung eines verbindenen Pfeils zwischen “ $\beta_{0j}$ “ und “ $\gamma_{00} + u_{0j}$ “ als Gleichung “ $\beta_{0j} = \gamma_{00} + u_{0j}$ “ gelesen werden muss:

$$\begin{array}{c} \beta_{0j} \\ \downarrow \\ = \gamma_{00} + u_{0j} \end{array}$$

Abb. 1f: Graphische Skizze des Übungsmodells (mit rein fiktiven Schätzwerten)



Die Abbildung 1f zeigt das zuvor abgedruckte Gleichungssystem unseres Übungsmodells mit den Gleichungen 1f bis 1k (vgl. die vorhergehende Seite). Die in Abbildung 1f zuoberst eingezeichnete L-Gleichung ist die L1-Gleichung (Gl. 1f), und die unterste L-Gleichung ist die L3-Gleichung (Gl. 1i). Dazwischen befindet sich die L2-Gleichung (1g). Die in der Abbildung enthaltenen Pfeile symbolisieren keine Kausalitäten, sondern verknüpfen immer zwei Gleichungsseiten miteinander. So bedeutet z.B. der senkrechte, durchgehende Pfeil zwischen

“ $\gamma_{00k}$ “ und “ $= \delta_{000} + \delta_{001} W_k + v_{00k}$ “ (auf der linken, unteren Seite der Abbildung 1f) die Gleichung:

$$\gamma_{00k} = \delta_{000} + \delta_{001} W_k + v_{00k} \text{ (Gl. 1i).}$$

Auf diese Art und Weise werden die drei Hauptgleichungen des Gleichungssystems (1f, 1g, 1i) graphisch miteinander verbunden. Die gestrichelten senkrechten Pfeile mit den Endpunkten “ $\delta_{010}$ ,  $\delta_{100}$ ,  $\gamma_{10k}$ “ bedeuten, dass diese Parameter ohne Variation zwischen Schulklassen und/oder Schulen als feste (fixed) Parameter geschätzt werden.<sup>18</sup> Die in Klammern angegebenen Schätz

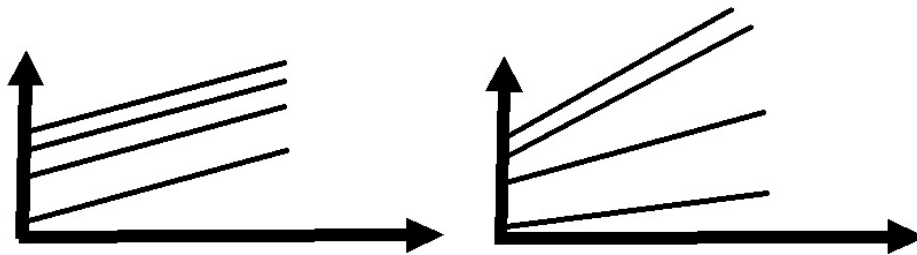
<sup>18</sup> In Mehrebenen-Modellen können auf der höchsten Ebene keine Random-Slopes bzw. Random-Effekte spezifiziert werden. Deshalb ist der Effekt “ $\delta_{001}$ “ auf der höchsten Ebene unseres Modells (die in Abb. 1f als unterste Ebene aufgenommen wurde) auch “fixed”, obwohl

werte “5,00“ und “500,00“ sind rein fiktive Werte und symbolisieren nur, dass an diesen Stellen die tatsächlich geschätzten Werte in die Grafik eingesetzt werden müssten. Die gepunktete, obere waagrechte Linie bedeutet, dass an dieser Stelle die Gleichung 1f allein zu Darstellungszwecken auseinander gerissen wurde und beide Teile außerhalb der Abbildung natürlich ohne Abstand zusammengeführt werden müssen:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{ijk} + e_{ijk} \quad (1f)$$

In allen unseren Überlegungen zur Konstruktionslogik unseres Übungsmodells haben wir bislang unterstellt, dass die jeweiligen Slopes (bzw. Steigungskoeffizienten, bzw. partiellen Regressionskoeffizienten, bzw. partiellen Effekte) über alle Personen, Schulklassen und Schulen konstant bleiben. Dass sie mithin “fixed effects” und nicht “random effects” sind. Die folgende Abbildung 1g macht den Unterschied zwischen diesen beiden Effekt-Typen deutlich.

Abb. 1g: “Fixed effects” versus “random effects”



Das linke Koordinatensystem zeigt vier Regressionsgeraden mit einer Kombination aus random Intercepts und fixed Slopes. D.h. die Intercepts sind unterschiedlich groß (random) und die Slopes sind konstant (fixed). Demgegenüber zeigt das rechte Koordinatensystem vier Regressionsgeraden mit einer Kombination aus random Intercepts und random Slopes. D.h. die Intercepts und auch die Slopes sind unterschiedlich groß. Letzteres bedeutet z.B., dass der Effekt von WEIBL\_KL auf  $b_0$  von Schulklasse zu Schulklasse unterschiedlich groß sein kann (random effect).

Im oben gezeigten System aus L1-, L2- und L3-Gleichung erzeugen bei einer als “random” definierten Modellschätzung die variierenden X-, Z- und/oder W-Effekte die Varianzkomponenten der Regressionsgleichungen. So entsteht die Varianz eines Slopes, wenn sich dieser über die Cluster des Mehrebenen-Modells hinweg verändert.

Ob X-, Z- und/oder W-Effekte im Modell als “random effects” geschätzt werden sollen, bestimmt der Analyst. Oftmals werden sowohl Modelle mit random Effekten als auch Modelle mit fixed Effekten geschätzt und werden deren Anpassungserfolge mittels Vergleich von AIC- und BIC-Werten (s.u.) untereinander verglichen.

Hinsichtlich der Schätzung von random Slopes bedeuten variierende Effekte im einfachsten Fall, dass ihre numerische Stärke durch Berechnung des Mittelwerts von allen Slopes ermittelt werden kann.

---

der Effekt nicht als Endpunkt einer gestrichelten Linie eingezeichnet wurde.

## B2 Was bedeuten die Konfidenzintervalle von Random-Effekten?

Wir wollen diese Frage am Beispiel von Modell 06 in diesem Skript beantworten (vgl. Kap. B4). Dort wird ein hochsignifikanter Random-Effekt " $\gamma_{10}=0,348$ " des Erlebens positiver Gefühle (X1-Variable: PA) auf die erlebte Lebenszufriedenheit (Y-Variable: LZ) geschätzt. Dieser Effekt variiert aber stark zwischen den Befragten in den 40 verschiedenen Ländern. So ist er unter den Befragten des einen Landes stark und unter den Befragten eines anderen Landes nur schwach ausgeprägt. Die Varianz dieses Effekts bzw. die diesbezügliche Fehlervarianz " $\text{Var}(u_{ij})$ " beträgt demnach über alle 40 Länder hinweg 0,006 und ist hoch signifikant.

Was bedeutet dieser Varianzwert von 0,006? Ist er vielleicht aufgrund seiner geringen Größe zu vernachlässigen und kann deshalb angenommen werden, dass die PA-Effekte in allen 40 Ländern eher gleich groß sind? Um dies zu überprüfen, müssen wir folgende Kalkulation vornehmen:

Wenn wir davon ausgehen können, dass (prinzipiell gedacht) sich sehr viele Schätzungen des Effekts analog zu einer Normalverteilung verteilen, und wenn für jede einzelne Schätzung sehr viele Beobachtungen (mindestens  $N=120$ ) zur Verfügung stehen, so können nach der folgenden Formel die Grenzen (Gr) eines 95%-Konfidenzintervalls berechnet werden:

$$\text{Gr}(\text{KI}_{0,95}) = \gamma_{10} \pm 1,96 \sqrt{[\text{Var}(u_{ij})]}$$

In dieser Formel bezeichnet " $\sqrt{[\text{Var}(u_{ij})]}$ " die Standardabweichung (SD) der Verteilung, die sich aus der Quadratwurzel der Varianz ergibt. Wir können also auch schreiben:

$$\text{Gr}(\text{KI}_{0,95}) = \gamma_{10} \pm 1,96 * \text{SD}$$

Mit unseren Daten (s.o.) ergeben sich daraus folgende Intervallgrenzen:

$$\text{Gr}(\text{KI}_{0,95}) = 0,348 \pm 1,96 * \sqrt{[0,006]}$$

$$\text{Gr}(\text{KI}_{0,95}) = 0,348 \pm 1,96 * 0,077$$

$$\text{Gr}(\text{KI}_{0,95}) = 0,348 \pm 0,150$$

$$\text{unterer Grenzwert: } 0,348 - 0,150 = 0,198$$

$$\text{oberer Grenzwert: } 0,348 + 0,150 = 0,498$$

Demnach liegen 95% der in allen 40 Ländern geschätzten PA-Effekte im Intervall zwischen 0,198 und 0,498. Anders beschrieben liegen die geschätzten Effekte mit einer Sicherheit von 95% zwischen den Werten von 0,198 und 0,498 (vgl. dazu auch Hox et al 2018: 16).<sup>19</sup>

(ACHTUNG: bei der Berechnung von Gr(KI) wird nicht der Standardfehler (SE) sondern die Standardabweichung (SD) benutzt.)

<sup>19</sup> Sollte die Schwankungsbreite zu groß ausfallen und kleiner gewünscht werden, kann versucht werden, dies durch Aufnahme zusätzlicher Kontextvariablen und/oder von cross-lagged Effekten zu erreichen.

### B3 Welche Bedeutung haben Kovarianzen zwischen den Fehlertermen?

Wenn in einem Mehrebenen-Modell mindestens zwei Komponenten auf der gleichen Ebene als “random” definiert werden, so können die Varianzen dieser Random-Komponenten auch miteinander kovariieren. Allerdings bedeutet diese Kovarianz nicht einfach nur, dass zwei Residuen-Varianzen miteinander zusammenhängen. Sie kann auch bedeuten, dass es einen Zusammenhang zwischen dem Random-Intercept des Modells und dem Random-Effekt einer unabhängigen Variablen geben kann.

Wir wollen dies am Beispiel von Modell 06 verdeutlichen und zeigen deshalb an dieser Stelle schon einmal die formale und auf der folgenden Seite die graphische Beschreibung des Modells (mehr dazu in Kap. B4).

$$\text{L1-Gleichung: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (6.1)$$

$$\text{L2-Gleichungen } \beta_{0j} = \gamma_{00} + u_{0j} \quad (6.2a)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (6.2b)$$

$$\text{“fixed part”}: \gamma_{00}, \gamma_{10}$$

$$\text{“random part”}: \text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$$

$$\text{“covariance”}: \text{Cov}(u_{0j}, u_{1j})$$

Das Modell 06 besteht aus zwei Ebenen mit einem Prädiktor “X” auf der Ebene L1. Die Ebene L1 wird mit der Gleichung 6.1 beschrieben. Hinzu kommt die Definition der Ebene 2 mit den beiden Gleichungen 6.2a und 6.2b. Auf der folgenden Seite wird das oben aufgeführte Gleichungssystem auch mit der Abbildung 6.1 graphisch veranschaulicht. Die darin enthaltenen Pfeile verknüpfen wie immer zwei Gleichungsseiten miteinander, sodass z.B. der langgezogene Pfeil in Abb. 6.1 (s.u.) gelesen werden muss als:  $\beta_{0j} = \gamma_{00} + u_{0j}$

$$\begin{array}{c} \beta_{0j} \\ \downarrow \\ = \gamma_{00} + u_{0j} \end{array}$$

Das Intercept von Gl. 6.1 variiert auf der L2-Ebene, was mit der Gl. 6.2a verdeutlicht wird. Je nach Kontext “j” kann es unterschiedlich groß ausfallen. Es setzt sich zusammen aus einer festen (fixed) Größe “ $\gamma_{00}$ ” und einem Residuum “ $u_{0j}$ “, das je nach Kontext “j” unterschiedlich groß ausfällt und deshalb auch die Varianz des Intercepts ausmacht (vgl. Gl. 6.2a).

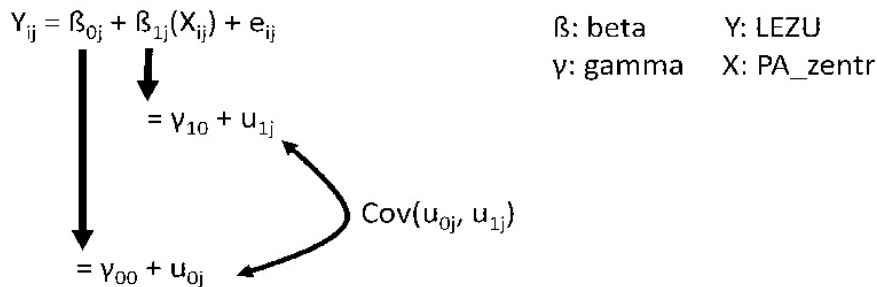
Ebenso verhält es sich mit dem Effekt bzw. dem Regressionskoeffizienten von Prädiktor “X”, der hier als “ $\beta_{1j}$ ” bezeichnet wird. Dieser Effekt bzw. dieser Slope variiert ebenfalls auf der Ebene L2 und nimmt je nach Kontext “j” unterschiedliche Größen an. Diese bestehen aus einem festen (fixed) Wert “ $\gamma_{10}$ ” und einem Residuum “ $u_{1j}$ “, das je nach Kontext “j” unterschiedliche Werte annimmt und deshalb die Varianz des Slopes erzeugt (vgl. Gl. 6.2b).

Die beiden Residuentерme “ $u_{0j}$ ” und “ $u_{1j}$ ” können miteinander kovariieren und diese Kovarianz ist auch schon in der folgenden Abbildung 6.1 als geschweifte Linie eingezeichnet. Diese Kovarianz bedeutet jetzt nicht nur, dass die Residuenvarianzen miteinander kovariieren, sondern auch, dass Intercept und Steigungskoeffizient miteinander kovariieren.

Was bedeutet das inhaltlich? Unsere abhängige Variable “Y” in Gl. 6.1 ist die gemessene Lebenszufriedenheit (LEZU) von Befragten in 40 Ländern . Unser Prädiktor X in Gl. 6.1 ist die zentrierte Messung einer positiv-affektuellen Weltsicht (PA), mit der die Befragten ihr Leben betrachten. So kann eine deutliche Kovarianz zwischen beiden Residuen auf das Folgende verweisen:

Die  $u_{0j}$ -Residuen bezeichnen die unterschiedlich großen Y-Achsenabschnitte bzw. die unterschiedlich groß ausgeprägte positiv-affektuelle Weltsicht (PA) der Befragten, weil sich diese nach Gl. 6.2a aus der Differenz zwischen  $\beta_{0j}$  und  $\gamma_{00}$  ergibt ( $u_{0j} = \beta_{0j} - \gamma_{00}$ ). Dementsprechend bezeichnen die  $u_{1j}$ -Residuen unterschiedlich stark ausgeprägte Regressionskoeffizienten “ $\beta_{1j}$ “ bzw. den unterschiedlich stark ausgeprägten Effekt von PA auf LEZU, denn dieser ergibt sich aus der Differenz zwischen  $\beta_{1j}$  und  $\gamma_{10}$  ( $u_{1j} = \beta_{1j} - \gamma_{10}$ ). Eine Kovarianz zwischen beiden Differenzen bedeutet somit: wenn in den verschiedenen Ländern die Lebenszufriedenheit (LEZU) groß ist, dann ist auch der Effekt von PA auf LEZU groß (oder: wenn das eine niedrig ist, ist auch das andere gering). Im Falle von negativen Kovarianzen wären gegenläufige Zusammenhänge zu beobachten gewesen (je stärker das eine, umso schwächer wäre das andere). In unserem Beispiel ist die betreffende Kovarianz allerdings positiv ausgeprägt und schwach signifikant ( $p=0,035$ ). Vgl. dazu die Erläuterungen zu Modell 06 in Kap. B4.

Abb. 6.1: Graphische Skizze des Gleichungssystems von Modell 06



Es gibt für MLAs viele verschiedene Typen von Kovarianz-Matrizen. Nur drei davon sollen im Folgenden kurz vorgestellt werden:

- 1) Wenn im Modell nur das Intercept variiert (Random-Intercept), kann es auch keine Kovarianzen geben. Dann wird auf Modellebene eine Identitätsmatrix bestimmt, die nur die Varianz von  $u_0$  enthält, sodass alle Zellen auf der Diagonalen gleich besetzt sind (Varianz) und alle anderen Zellen als Null-Zellen zu betrachten sind. Dieser Matrix-Typ wird für “Random Intercept only-Modelle” benutzt (wie z.B. für Modell 01):

$$\text{Var}(u_{0j}) \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

- 2) Wenn es im Modell ein Random-Intercept und einen Random-Slope gibt (wie z.B. für das oben erläuterte Modell 06), werden auch beide Varianzen aufweisen und diese Varianzen

können auch miteinander kovariieren. In Mplus (vgl. Kap. C) werden jedoch zunächst einmal (als Voreinstellung) keine Kovarianzen geschätzt. Dementsprechend sind zunächst einmal in Mplus die MLA-Modelle durch eine diagonale Kovarianzmatrix gekennzeichnet, deren Varianzen (in der Diagonalen) unterschiedlich groß sein können:

$$\begin{vmatrix} \text{Var}(u_{0j}) & 0 \\ 0 & \text{Var}(u_{1j}) \end{vmatrix}$$

- 3) In den meisten, der im Folgenden vorgestellten MLAs werden wir sowohl die Varianzen der Residuen (bzw. der Random-Intercepts und Random-Effekte) als auch die Kovarianzen zwischen den Residuen (bzw. zwischen den Intercepts und Random Effekten) schätzen. Von daher werden wir für jedes unserer Modelle eine unstrukturierte Kovarianzmatrix voraussetzen, in der alle Matrixelemente unterschiedlich groß sein können und in der alle Matrix-Zellen mit diesen Elementen besetzt werden können. Die Kovarianzstruktur der Residuen (“within” und “between”) sollte offen sein für alle möglichen Varianzen und Kovarianzen der Residuen. Insbesondere dann, wenn keine diesbezüglichen Hypothesen formuliert werden können (was wohl die Regel sein dürfte). Zudem in der Forschungspraxis die Wahl eines bestimmten Kovarianztyps nicht besonders relevant für die Ermittlung von Parameterschätzwerten und Standardfehlern ist (nach: Bickel 2007: 92).

Die unstrukturierte Matrix der Varianzen und Kovarianzen aller Residuen ist demnach von folgender Form:

$$\begin{vmatrix} \text{Var}(u_{0j}) & \text{Cov}(u_{1j} u_{0j}) \\ \text{Cov}(u_{1j} u_{0j}) & \text{Var}(u_{1j}) \end{vmatrix}$$

Andere, hier nicht behandelte Matrix-Typen sind z.B. die AR(1)-Matrizen, die eine first order autoregressive Struktur aufweisen. Sie werden vor allem für Panel-Studien benutzt (die im vorliegenden Skript nicht behandelt werden) und implizieren, dass je näher die Kontexte beieinander liegen (z.B. “1 dann 2” versus “1 dann 3”), umso größer die Kovarianz sein sollte.

Wenn es in einer MLA keine Möglichkeit gibt, den Typ von Kovarianz-Matrix bereits vor Beginn der Modellschätzung festzulegen, können verschiedene Modelle mit unterschiedlichen Matrix-Typen geschätzt werden und die Modellschätzungen mittels AIC/BIC (vgl. Kap. B5) miteinander verglichen werden. Wir werden darauf in diesem Skript nicht weiter eingehen.

Wenn der für eine MLA spezifizierte Matrix-Typ zu wenig komplex ist (also z.B. zu wenige Residuen-Kovarianzen aufweist), unterliegen die Parameter-Schätzwerte eher einem Typ 1-Fehler, d.h. die Parameter-Schätzwerte sind dann eher falsch signifikant. Bei einer zu hohen Komplexität des spezifizierten Matrix-Typs (also z.B. bei zu vielen spezifizierten Residuen-Kovarianzen) kann es eher zu einem Typ 2-Fehler kommen, d.h. die Parameter-Schätzwerte können dann eher falsch nicht-signifikant werden (nach: Field et al. 2012: 860-872).



#### **B4 Wann und wie sollte zentriert werden?**

Zur Beseitigung von Interpretationsproblemen werden die X-Variablen in einer MLA häufig zentriert (die Y-Variable wird nicht zentriert!). Wenn es z.B.:

- (-) negativ geschätzte Interceptwerte gibt, die aufgrund der Y-Skala eigentlich nicht vorkommen dürften,
  - (-) oder wenn bei den geschätzten Intercepts überraschend Nullwerte auftauchen, die auch nicht zulässig wären,
  - (-) oder wenn die Prädiktoren keinen empirischen Nullwert haben,
- dann lassen sich die damit verbundenen Interpretationsprobleme dadurch beseitigen, dass die X-Variablen vor der Schätzung zentriert werden.

Bei zentrierten Variablen bekommt der Nullwert eine Bedeutung, denn nach der Zentrierung entspricht der Nullwert dem Durchschnittswert einer Verteilung. Negative Werte sind dann unterdurchschnittliche Ausprägungen und positive Werte überdurchschnittliche Ausprägungen einer Variablen. Ein geschätzter Interceptwert wäre dann so zu interpretieren, dass dieser Wert zu erwarten ist, wenn alle X-Variablenwerte auf 0 gesetzt werden bzw. ihren jeweiligen Durchschnittswert aufweisen. Ein "Nullwert" bei den Prädiktoren bedeutet nach der Zentrierung "Durchschnittswert".

Dichotome X bzw. Z-Prädiktoren (wie z.B. "SEX2" 0/1, oder der nach Mediansplit entstehende Prädiktor "PA2") können zwar, rein formal-statistisch betrachtet, zentriert werden, sollten es aber nicht, weil die Effekte der Dichotomie "0/1" einfach und empirisch eindeutig zu interpretieren sind. Sie sollten auch nicht standardisiert werden.

In der MLA werden zwei unterschiedliche Zentrierungsformen benutzt:

- a) die Zentrierung am Grand Mean (= Mittelwert aller Beobachtungswerte unabhängig von deren Clusterzugehörigkeiten),
- b) die Zentrierung am Mittelwert des jeweiligen Clusters bzw. der jeweiligen Gruppe (Gruppenmittelwert).

Üblicherweise wird in einer MLA bei Zentrierung von metrischen Variablen die Zentrierung am Grand Mean (a) gewählt. Allgemein lässt sich sagen, dass eine Zentrierung mit dem Grand Mean gewählt werden sollte, wenn das Forschungsinteresse auf der L2-Ebene liegt. Demgegenüber sollte eine gruppenspezifische Zentrierung gewählt werden, wenn das Forschungsinteresse auf der L1-Ebene liegt oder wenn das Interesse auf die Interpretation von cross-level Effekten ausgerichtet ist.

Die Grand-Mean-Zentrierung analysiert die einzelnen Messwerte in Bezug auf das ganze Sample und ignoriert gruppenspezifische Unterschiede, während die gruppenbezogene Zentrierung die Messwerte in den Kontext einer jeweiligen Gruppe setzt. So kann ein zentrierter Messwert bei Grand-Mean-Zentrierung überdurchschnittlich und bei gruppenbezogener Zentrierung unterdurchschnittlich sein. Die Entscheidung zwischen den beiden Verfahren sollte somit eher analytisch und nicht formal-statistisch getroffen werden.

Die Schätzergebnisse bei Zentrierung am Grand Mean (a) und bei Zentrierung am Gruppen-Mean (b) sind sich eher ähnlich und oftmals sind die dabei entstehenden Unterschiede ohne substantielle Bedeutung. Wenn es Random Intercepts gibt, ist auch der Modellfit einer MLA-Schätzung bei a) und b) identisch. Unterschiede gibt es vor allem bei den geschätzten L2-Varianzen (Heck/Thomas 2015: 87-93).

Eine weitere Form der Zentrierung ist die Standardisierung von Variablen, bei welcher deren Mittelwert auf null und deren Standardabweichung auf 1 gesetzt wird (vgl. auch Kap. C3). Dadurch verändert sich aber auch die Metrik der standardisierten Variablen, die dann in SD-Einheiten ausgedrückt wird und ein Regressionsgewicht von 3,75 würde z.B. bedeuten, dass bei Anwachsen des Prädiktors um eine Standardeinheit die abhängige Variable um 3,75 empirische Einheiten anwächst.

Die oftmals vorgetragene Argumentation, wonach X-Variablen, die hoch miteinander korrelieren und somit Multikollinearitäten erzeugen, durch Zentrierung aus ihren gegenseitigen linearen Abhängigkeiten befreit werden könnten (vgl. z.B. Finch/Bolin 2017: 38f) ist irreführend und auch mathematisch unbegründet (vgl. Urban/Mayerl 2018: 248-252).

## **B5 Wie sollten Modelle miteinander verglichen werden? (AIC, BIC, $R^2$ )**

In der MLA-Praxis sollen häufig zwei oder mehrere Modelle miteinander verglichen werden, um festzustellen, welches Modell das “beste” ist. Üblicherweise werden dazu statistische Modelle mittels eines Chi-Quadrat-Differenzentests untereinander verglichen. Dies setzt aber voraus, dass die Modelle<sup>20</sup>

- a) mit einem “reinen” Maximum-Likelihood-Schätzverfahren (z.B. mit FML oder RML<sup>21</sup>) berechnet wurden,
- b) hierarchisch geschachtelt (nested) sind,
- c) die gleiche abhängige Variable aufweisen,
- d) mit gleichen Fallzahlen geschätzt wurden.

Bei den im vorliegenden Skript geschätzten Modellen wird ein robustes ML-Verfahren eingesetzt (vgl. Kap. A2, Pkt. 3) und ist folglich die Bedingung (a) nicht erfüllt. Zudem können unsere Modelle oftmals nicht als “geschachtelt” betrachtet werden, weil sie oftmals unterschiedliche Variablen und unterschiedliche Analyseebenen besitzen und deshalb gegen Bedingung (b) verstoßen können. Allerdings haben alle Modelle mit Variablen aus nur einem Pool (Modell 01 bis 14, oder Modell 20 bis 28) die gleiche abh. Variable und würden somit Bedingung (c) erfüllen. Auch Bedingung (d) wäre bei Vergleichen von Modellen aus dem Pool der Modelle 01 bis 14 sowie aus dem Pool der Modelle 20 bis 28 erfüllt: Alle Modelle, die bei uns im gleichen Modell-Pool enthalten sind, werden mit der gleichen Anzahl von Fällen (auf allen Ebenen) geschätzt.

Für einen Modellvergleich können wir somit keinen Chi-Quadrat-Differenzentest vorschlagen, sondern empfehlen dazu einen Vergleich der modellspezifischen AIC- bzw. BIC-Werte. Diese Index-Werte werden von uns auch bei allen Modellvergleichen im vorliegenden Skript benutzt und interpretiert.

Die AIC- und BIC-Indizes sind Informationsmaße. Sie berücksichtigen den Fit und die Komplexität von Modellen und können auch eingesetzt werden, wenn die zu vergleichenden Modelle mit robusten ML-Schätzverfahren (MLR) berechnet wurden und/oder nicht-geschachtelt sind.

---

<sup>20</sup> Vgl. dazu Urban/Mayerl 2014: 218-227.

<sup>21</sup> Bei FML-Schätzungen (full maximum likelihood-Schätzungen) werden alle zu schätzenden Parameter (Regressionskoeffizienten und Varianzkomponenten) gleichzeitig in die iterative Schätzung einbezogen, während bei RML-Schätzungen (restricted maximum likelihood-Schätzungen) zuerst die Schätzung der Varianzkomponenten und dann erst die Schätzung der Regressionskoeffizienten erfolgt (nach Poetschke 2014: 227).

Nach Hox et al. 2018: 39 können bei FML-Schätzung allein solche Modellschätzungen miteinander verglichen werden, die sich nur entweder im “fixed part” oder im “random part” unterscheiden, während sich bei RML-Schätzung die Modelle nur im “random part” unterscheiden dürften. Diese Forderungen wurden allerdings bislang in der Literatur nicht diskutiert und somit auch nicht akzeptiert oder verworfen.

Je kleiner ein AIC/BIC-Wert ist, umso besser ist die Modellschätzung. Modellschätzungen mit AIC/BIC-Werten, die kleiner als die Werte von anderen Modellschätzungen sind, sind besser als Modellschätzungen mit höheren Werten. Dabei sollten Differenzen bei den AIC/BIC-Werten nicht metrisch, sondern nur ordinal interpretiert werden (“welches Modell ist besser”, “welches Modell ist deutlich besser” o.ä.). Denn die AIC/BIC-Werte haben keine intrinsische Bedeutung und sind deshalb auch nicht intrinsisch interpretierbar.

So wird z.B. eine Differenz zwischen den BIC-Werten von zwei Modellen, die bei 6 bis 10 Punkten liegt, als Hinweis auf einen deutlichen Unterschied zwischen den beiden Modellen gewertet, sodass das Modell mit dem kleineren BIC-Wert als “besseres Modell” bevorzugt werden kann. Und Differenzen zwischen den BIC-Werten von mehr als 10 Punkten verweisen auf das Modell mit der niedrigeren Punktezahl als “sehr viel besseres Modell”.<sup>22</sup>

Dementsprechend sind auch die Differenzen bei AIC-Werten zu interpretieren. Allerdings hat sich hierbei in der Forschungspraxis durchgesetzt, dass schon AIC-Werte, die um zwei Punkte niedriger sind, als Hinweis auf ein deutlich besseres Modell interpretiert werden.

Der AIC-Index (Akaike’s information criterion) und der BIC-Index (Bayesian information criterium) werden berechnet nach:

$$\text{AIC} = (-2 * \ln L) + (2 * q)$$

$$\text{BIC} = (-2 * \ln L) + (2 * q * \ln(N))$$

wobei bedeuten:

lnL: Minimum der Log-Likelihood-Funktion bei der Modellschätzung

q: Anzahl freier, zu schätzender Modellparameter

ln(N): natürlicher Logarithmus der Anzahl von Beobachtungsfällen (auch als “H<sub>0</sub>-value” bezeichnet), wobei es unklar ist, welches N von welcher L-Ebene benutzt werden sollte.

Aufgrund der Ungewissheit bei der Berechnung von ln(N) (s.o.) sollte in der Mehrebenen-Analyse eher der AIC-Index oder sollten zumindest beide Indexwerte berichtet werden.<sup>23</sup> Oftmals wird auch empfohlen, bei MLAs mit großen Clustern und wenigen Parametern den BIC für einen Modellvergleich zu benutzen.

Eine spezielle Form des Modellvergleichs in der Mehrebenen-Analyse wird durch die Berechnung eines Pseudo-R<sup>2</sup> ermöglicht. Die Logik des aus der OLS-Regression bekannten Determinationskoeffizienten “R<sup>2</sup>” wird dabei auch für die Mehrebenen-Analyse genutzt, indem die Residualvarianz des Null-Modells (hier: Modell 01) mit der Residualvarianz eines zu testenden Mehrebenen-Modells (im Folgenden: Modell 03) verglichen wird. Ist die Residualvarianz des zu testenden Modells geringer als die Residualvarianz des Nullmodells, so kann dies als Erfolg von dessen Modellspezifikation gewertet werden. Dann kann das prozentuale Ausmaß dieser Va-

---

<sup>22</sup> Vgl. dazu die Diskussion auf: [www.statmodel.com/discussion/messages/11/5682.html](http://www.statmodel.com/discussion/messages/11/5682.html).

<sup>23</sup> In der MLA mit Mplus wird ein “sample-size adjusted BIC” berichtet, das im Folgenden auch bei den Mplus-Output-Files ausgedruckt wird, das aber in der Literatur noch nicht diskutiert wurde.

rianzreduktion als Wert eines Pseudo-R<sup>2</sup>-Koeffizienten interpretiert werden. Das Pseudo-R<sup>2</sup> berechnet sich als:

$$\text{Pseudo-R}^2 = \frac{(\text{Summe der Resid.varianz, Nullmodell}) - (\text{Summe der Resid.varianz, Testmodell})}{(\text{Summe der Resid.varianz, Nullmodell})}$$

Verdeutlicht sei diese Formel anhand der Schätzergebnisse unserer Modelle 01 (Nullmodell ohne Prädiktoren) und 03 (Testmodell mit Prädiktoren). Die Residualvarianz des Modell 01 (Nullmodell, keine Prädiktoren) ergibt sich aus dessen Within-Residualvarianz (1,229) plus dessen Between-Residualvarianz (0,279) und beträgt somit 1,508. Die Residualvarianz des Modell 03 (Testmodell, mit Prädiktoren) ergibt sich aus dessen Within-Residualvarianz (1,003) plus dessen Between-Residualvarianz (0,179) und beträgt somit 1,182. Werden diese Werte in die obige Formel eingesetzt, berechnet sich daraus ein Pseudo-R<sup>2</sup> von 0,216. Mithin hat die Spezifikation und Schätzung von Modell 03 die Residualvarianz von Modell 01 um 21,6% reduziert, hat also die Schätzung der abh. Variablen LEZU um einen deutlichen Betrag verbessert (sollten im Testmodell auch Random Slopes geschätzt worden sein, kann deren Varianz zur nächsthöheren Ebene hinzu addiert werden).

Im obigen Beispiel wurde das Pseudo-R<sup>2</sup> für das gesamte Modells 03 (bestehend aus einer Ebenendifferenzierung und zwei L1-Prädiktoren) ermittelt. In der Forschung wird auch häufig die Varianzreduktion nur auf einer Ebene des Testmodells berechnet.

Die Kalkulation eines Pseudo-R<sup>2</sup> ist in der Mehrebenen-Analyse nicht ohne Probleme. Die drei wichtigsten sind: a) wenn ein Prädiktor die erklärte Varianz reduziert, kann u.U. ein negatives Pseudo-R<sup>2</sup> berechnet werden, b) für Modelle mit Random Slopes hängt die geschätzte Varianz von den Skalierungen der Prädiktoren ab, c) bei Modellen mit Cross-Level-Effekten muss die Residualvarianz deutlich komplexer berechnet werden (vgl. dazu Hox et al. 2018: 57-64). In diesen Fällen sollte das Pseudo-R<sup>2</sup> nur sehr vorsichtig (wenn überhaupt) benutzt werden.

Eine Alternative zum Pseudo-R<sup>2</sup> ist das Maddala-R<sup>2</sup>, das aus der Differenz der Devianzen von Nullmodell-Schätzung und Testmodell-Schätzung berechnet wird (nach: Poetschke 2006: 174):

$$\text{Maddala-R}^2 = 1 - \exp[ (\text{Devianz, Testmodell} - \text{Devianz, Nullmodell}) / (\text{Fallzahl}) ]$$

Das dabei üblicherweise benutzte Devianzmaß ist “-2loglikelihood” (-2LL). Im Mplus-Output ist “LL” als “Loglikelihood, H0 Value” direkt zu Beginn der “MODEL FIT INFORMATION” zu finden. Wir wollen die Berechnung von Maddala-R<sup>2</sup> am Beispiel von Modell 03 verdeutlichen:

$$-2LL(\text{Nullmodell}) = -2(-10616,565) = 21233,13$$

$$-2LL(\text{Testmodell}) = -2(-9910,468) = 19820,936$$

$$[ (19820,936 - 21233,13) / (6929) ] = -1412,194 / 6929 = -0,203$$

$$1 - \exp[ -0,203 ] = 1 - 0,81 = 0,19$$

Somit beträgt der Anteil an Varianz im Nullmodell, der mit Modell 03 ausgeschöpft wird, nach Berechnung von Maddala- $R^2$  insgesamt 19% und ist damit im Vergleich zur Berechnung nach Pseudo- $R^2$  leicht geringer (19,0% statt 21,6%).

## C Mehrebenen-Analyse mit Mplus

### C1 Modellierung und Analysestrategie

Die Konstruktion, Schätzung und Analyse von Mehrebenen-Modellen wird im vorliegenden Skript anhand von 23 Beispielen erläutert. Diese Beispiele zeigen Grundmuster der MLA unter Verwendung von Mplus und können im Kontext der Logik von MLA/Mplus beliebig modifiziert werden.

Grundsätzlich betrachtet gibt es zwei unterschiedliche Strategien zur Durchführung von Mehrebenen-Analysen: die “Top-Down”- und die “Bottom-Up”-Strategie.

Bei der “Top-Down”-Strategie beginnt die Analyse mit der analytischen Entwicklung und/oder theorie-orientierten Begründung eines Modells, das möglichst alle spezifizierten Effekte bereits auf der ersten Stufe der Analyse enthält. Wenn sich dann im Laufe der Analysen bestimmte Effekte als nicht bedeutsam herausstellen, kann das entsprechende Modell reduziert oder modifiziert werden. Es kann sodann die Schätzung des ursprünglichen Modells mit der Schätzung des modifizierten Modells verglichen und dasjenige Modell mit dem relativ besten Fit für die weitere Analyse beibehalten werden (vgl. dazu Kap. B5).

So wird z.B. in Modell 07 (vgl. Kap. D4) getestet, ob die positive Affekt-Orientierung (PA) von befragten College-Studierenden eine stärkere oder schwächere Bedeutung für deren Lebenszufriedenheit (LEZU) hat als deren männliche oder weibliche Geschlechtsausprägung (SEX2). In der empirischen MLA stellt sich heraus, dass beide unabh. Prädiktoren einen signifikant positiven Effekt auf die Höhe der Lebenszufriedenheit ausüben, dass aber nur der PA-Effekt auch signifikant zwischen den 40 Staaten, in denen die Befragung durchgeführt wurde, variiert. Es kann jedoch mit der Schätzung von Modell 07 nicht ermittelt werden, ob beide Effekte gleich stark sind, oder ob ein Effekt stärker als der andere Effekt ist. Dies kann mittels der Schätzung von Modell 07 nicht bestimmt werden, weil die Skalen beider Prädiktoren sehr unterschiedlich sind. Deshalb wird für die Schätzung des neu spezifizierten Modells 08 zwar die Modellstruktur von Modell 07 beibehalten, jedoch die Variable “positive Affekt-Orientierung” (PA2) dichotomisiert (in Form eines Mediansplits), sodass die Skalen beider Prädiktoren nunmehr dichotom sind und ihre Effekte dementsprechend direkt miteinander verglichen werden können. Nach der Schätzung von Modell 08 ist der dichotomisierte PA2-Effekt fast 6-mal so stark wie der dichotomisierte SEX2-Effekt, wobei die länderspezifische Variation des SEX2-Effekts auch in der Schätzung von Modell 08 nicht-signifikant ist.

Im hier berichteten Beispiel einer “TOP-Down”-Analyse werden also die Effekte von zwei Prädiktoren auf die Lebenszufriedenheit als linear, direkt und unabhängig voneinander sowie als länderspezifisch variierend festgelegt und werden diese Effekte sodann mittels zweier MLA untersucht.

In der “Bottom-Up”-Strategie beginnt die MLA in aller Regel nicht mit bestimmten Annahmen über makrospezifisch variierende Mikroeffekte, sondern folgt einem allgemein gültigen Muster von Mehrebenen-Analyse. Dabei durchläuft die Analyse die folgenden drei bis vier Stufen:

Stufe 1: Die MLA beginnt mit einem “leeren” Modell ohne Prädiktoren, bei dem nur die Intercepts zwischen den Makroeinheiten (in Modell 01: 40 Länder/Staaten) variieren (vgl. Kap. D4). So wird ein durchschnittlich gültiges Niveau der abh. Variablen (in Modell 01: LEZU) gefunden. In unserem Modellbeispiel 01 ist dies ein LEZU-Wert von 4,39 auf der LEZU-Skala von 1,00 bis 7,00. Da für dieses Beispiel ein ICC-Wert von 0,18 berechnet werden kann und demnach ca. 18% der Gesamtvarianz von LEZU durch Unterschiede zwischen den Makroeinheiten entstehen, indiziert die Modellschätzung von Modell 01 auch, dass es sinnvoll ist, die LEZU-Daten mittels MLA zu untersuchen.

Stufe 2: Nunmehr wird das Modellbeispiel (“random intercept model”) um die beiden Prädiktoren PA und SEX2 erweitert, von denen fixierte (“fixed”) Effekte, also Effekte, die nicht zwischen den verschiedenen Ländern variieren, ausgehen. Ein solches Modell entspricht unserem Beispielsmodell 03 (vgl. Kap. D4). Um zu ermitteln, wieviel Prozent der LEZU-Makrounterschiede durch die konstanten Effekte von PA und SEX2 gebunden werden, kann folgender Quotient berechnet werden:

$$1 - (\text{Between Varianz, Modell 03} / \text{Between Varianz, Modell 01}) = 0,358$$

Dementsprechend werden ca. 36% der LEZU-Makrounterschiede von Effekten der beiden Prädiktoren ausgeschöpft.

Stufe 3: Auf dieser Stufe werden die Effekte der beiden Prädiktoren zur länderspezifischen Schätzung freigegeben, d.h. es wird ein “random slopes”-Modell spezifiziert und geschätzt. In diesem Modell werden neben den Intercepts auch die Steigungen der PA- und SEX2-Effekte als variabel angesehen, sodass diese je nach Land, aus dem die Befragten kommen, unterschiedlich ausfallen können. Eine solche Modellspezifikation und Modellschätzung haben wir im Beispielsmodell 07 vorgenommen (vgl. Kap. D4). Ein Vergleich der AIC/BIC-Werte (vgl. Kap. B5) aus Modell 01 mit denjenigen aus Modell 07 zeigt, dass diese Werte sehr deutlich gefallen sind (z.B. der AIC-Wert von 21239,129 in M01 auf 19790,669 in M07) und somit mit Beispielsmodell 07 eine deutliche Verbesserung des Modellfits erreicht werden kann.

Stufe 4: Auf dieser Stufe der Modellanalyse werden auch cross-level-Effekte in die Modellschätzung mit einbezogen. Eine dementsprechende Modellspezifikation zeigt unser Beispielsmodell 12 (vgl. Kap. D4).<sup>24</sup> Die Modellschätzung zeigt allerdings, dass die beiden cross-level-Effekte von GDP2 (dichotomisiertes Bruttoinlandsprodukt) auf die Effekte von PA2 und SEX2 nicht signifikant sind und allein der Effekt von GDP2 auf die LEZU-Intercepts hochgradig signifikant und damit auch inhaltlich bedeutsam ist.

Eine weitere, häufig genutzte Methode zur Modifikation von Mehrebenen-Modellen besteht darin, mittels Aggregationsverfahren (häufig: durch Berechnung eines Mittelwerts) aus einem Prädiktor gleich mehrere zusätzliche Prädiktoren zu erzeugen. So kann z.B. in einem Drei-Ebenen-Modell aus einer Mikro-Variablen X eine zusätzliche Makro-Variable Z sowie eine

---

<sup>24</sup> Zu beachten ist, dass in Modell 12 alle Prädiktoren dichotomisiert wurden.



weitere Makro-Variable W und ein Teil einer cross-level-Interaktionsvariablen gewonnen werden. Wie ist das möglich?

Zur Verdeutlichung können wir die Variablen des hier in Kapitel E2 präsentierten Datensatzes "ch4three.DAT" benutzen. In diesem Datensatz gibt es die Variable:

X1 = SES (sozioökonomischer Status) für jede Herkunftsfamilie der befragten Schüler:innen (Analyseebene L1).

Aus X1 kann durch Berechnung eines SES-Mittelwerts für jede Schulklasse ein Durchschnitts-SES-Wert gebildet werden:

Z1 = durchschnittlicher SES (sozioökonomischer Status) einer jeden Schulklasse der befragten Schüler:innen (Analyseebene L2).

Aus Z1 kann durch Berechnung eines SES-Mittelwerts über alle SES-Mittelwerte aller Schulklassen ein Durchschnitts-SES-Wert für jede Schule ermittelt werden:

W1 = durchschnittlicher SES (sozioökonomischer Status) für jede Schule der befragten Schüler:innen (Analyseebene L3).

X\*Z/W = Jede Z- oder W-Variable kann als Teil einer cross-level Interaktionsvariablen benutzt werden. Diese informiert darüber, ob ein Individualeffekt in Abhängigkeit von Ausprägungen eines untersuchten Kontexts variiert. So kann z.B. die Bedeutung des "Geschlechts=weiblich" (L1) für die Mathematikleistung von befragten Schüler:innen (L1) in Abhängigkeit vom klassendurchschnittlichen SES-Wert (L2) variieren.

## C2 Mplus-Syntax und Mehrebenen-Analyse

Im Folgenden wollen wir einige grundsätzliche Hinweise zur Verwendung von Mplus für die Analyse von Mehrebenen-Modellen geben. Dabei ist zu berücksichtigen, dass wir in diesem Skript die Mplus-Syntax eher beispielsorientiert vermitteln wollen und deshalb an dieser Stelle nicht alle Mplus-Regeln zur Mehrebenen-Analyse vorgestellt werden.

Der Anwender sollte immer so vorgehen, dass er zunächst dasjenige Beispielsmodell von den 23 hier vorgestellten Modellen auswählt, das seinem Modell am ehesten entspricht. Dieses Modell kann dann modifiziert werden (z.B. durch zusätzliche Prädiktoren), wobei die im Folgenden erläuterten Syntax-Regeln hilfreich sein können.

In unseren Hinweisen werden folgende Variablen benutzt:

- C1 Cluster-ID für Ebene 1 (z.B. Personen bzw. Studierende)
- C2 Cluster-ID für Ebene 2 (z.B. Staaten/Länder oder Schulklassen)
- Y abh. L1-Variable (z.B. Lebenszufriedenheit)
- X1 L1-Prädiktor (z.B. Geschlecht, dichotom gemessen)
- X2 L1-Prädiktor (z.B. sozio-ökonomischer Status = SES)
- X3 L1-Prädiktor (z.B. Lebensalter)
- Z L2-Prädiktor (z.B. Bruttoinlandsprodukt der einzelnen Staaten/Länder aus C2 )

Grundsätzlich betrachtet kann in der Mehrebenen-Analyse der Effekt eines jeden Prädiktors spezifiziert werden:

- (A) als fester (fixed) Effekt/Slope,
- (B) als freier (random) Effekt/Slope,
- (C) als Teil eines cross-level Effekts/Slope.<sup>25</sup>

Diese drei unterschiedlichen Effekte können wir auch im Gleichungssystem eines Mehrebenen-Modells verdeutlichen. So zeigen wir auf der folgenden Seite das Gleichungssystem eines Zwei-Ebenen-Modells:

Auf der untersten Ebene (L1) (bzw. in der Abb. C2.1 auf der obersten Zeile) befindet sich die Lineargleichung (C2.1a) mit der abh. L1-Variablen "Y" und den drei L1-Prädiktoren "X1, X2 und X3". Diese drei Prädiktoren beeinflussen Y entsprechend der Stärke und Richtung der drei Gewichtungsfaktoren " $\beta_{1j}$ ,  $\beta_{2j}$  und  $\beta_{3j}$ ". Die Gewichtungsfaktoren variieren jedoch auch noch in unterschiedlicher Weise hinsichtlich des Kontexts, in dem sich die befragten Personen aufhalten (z.B. hinsichtlich des jeweiligen Landes/Staates). So ist  $\beta_{1j}$  in linearer Weise abhängig von Z ( $\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$ ). Und da Z eine L2-Variable ist, handelt es sich bei dem Effekt " $\gamma_{11}$ " um einen

---

<sup>25</sup> Dies gilt allerdings nicht für die Prädiktoren auf der höchsten Cluster-Stufe (hier: Z). Deren Effekte können nicht als Random-Effekte geschätzt werden. Allerdings können sie als Teil von cross-level Effekten oder zur Schätzung der Y-Intercepts eingesetzt werden (vgl. dazu Modell 12).

cross-level-Effekt bzw. um einen Effekt vom Typus (C). Die Gleichung C2.1c beschreibt also die Wirkungsweise eines cross-level-Effekts.

Im Unterschied dazu beschreibt die Gleichung C2.1d die Wirkungsweise des Effekts “ $\beta_{2j}$ “. Dieser ist ein Random-Effekt, der auf der L2-Ebene zwischen den verschiedenen L2-Einheiten (z.B. Länder/Staaten) variiert ( $\beta_{2j} = \gamma_{20} + u_{2j}$ ). Der Random-Effekt ist dementsprechend vom Typus (B). Der Effekt setzt sich zusammen aus einem festen Parameter “ $\gamma_{20}$ “ und den L2-spezifischen Residualgrößen “ $u_{2j}$ ”.

Anders verhält es sich mit dem Effekt von X3, der mit der Gleichung “ $\beta_{3j} = \gamma_{30}$ “ beschrieben wird (Gl. C2.1e). Er beeinflusst Y mit dem festen Wert von “ $\gamma_{30}$ “, der nicht in Abhängigkeit von L2-Einheiten variiert. Dementsprechend handelt es sich bei diesem Effekt um einen “fixed effect” oder um einen Effekt vom Typus (A).

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + \beta_{3j}X3_{ij} + e_{ij}$  (C2.1a)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (C2.1b)

$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$  (C2.1c)

$\beta_{2j} = \gamma_{20} + u_{2j}$  (C2.1d)

$\beta_{3j} = \gamma_{30}$  (C2.1e)

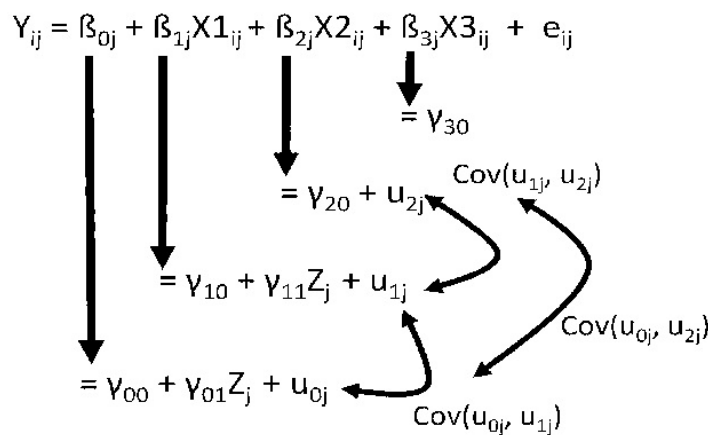
“fixed part”:  $\gamma_{00}, \gamma_{10}, \gamma_{20}, \gamma_{30}, \gamma_{01}, \gamma_{11}$

“random part”:  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j}), \text{Var}(u_{2j})$

“covariance”:  $\text{Cov}(u_{0j}, u_{1j}), \text{Cov}(u_{0j}, u_{2j}), \text{Cov}(u_{1j}, u_{2j})$

Da das oben beschriebene Gleichungssystem recht abstrakt und schwierig zu verstehen ist, wollen wir es noch graphisch veranschaulichen. Dies geschieht in der folgenden Abbildung C2.1. Wichtig zum Verständnis der darin enthaltenen Pfeile ist wiederum die Vorstellung, dass es sich bei den Pfeilen nicht um Einflusspfade, sondern um Verknüpfungen von jeweils zwei Gleichungsseiten handelt (vgl. dazu Kap. B3).

Abb. C2.1: Grafische Veranschaulichung des Gleichungssystems C2.1



Das Gleichungssystem C2.1 besteht aus fünf Gleichungen. (und drei Kovarianzen, die uns hier aber nicht weiter interessieren sollen). Das sind die Gleichungen C2.1a bis C2.1e. Vier dieser fünf Gleichungen werden in Abbildung C2.1 mit Pfeilen veranschaulicht. Der kürzeste Pfeil betrifft die Gleichung C2.1e und veranschaulicht einen “fixed effect”. Die mittleren beiden Pfeile

beziehen sich auf eine Gleichung mit einem Random-Effekt (dritter Pfeil von links) und auf eine Gleichung mit einem cross-level-Effekt (zweiter Pfeil von links). Der längste Pfeil (auf der äußerst linken Seite) symbolisiert die Abhängigkeit der Intercepts von Z und soll hier nicht weiter erläutert werden.

Wir wollen die Mplus-Spezifikationen der drei oben genannten Effekte in der folgenden Tabelle C2.1 verdeutlichen. Diese Spezifikationen sind Ausschnitte aus kompletten Mplus-Inputfiles mit noch weiteren, zusätzlichen Mplus-Anweisungen. Diese Mplus-Inputfiles sind die Modelle 03, 07 und 12. Sie sind im Folgenden in Kapitel D4 zu finden.

Tab. C2.1: Ausgewählte Mplus-Befehle für die Effekt-Typen A, B und C

Beispiel (A):	Beispiel (B):	Beispiel (C):
Zwei-Ebenen-Modell mit Random-Intercepts mit fixed Slopes von X1, X2	Zwei-Ebenen-Modell mit Random-Intercepts mit Random-Slopes von X1, X2	Zwei-Ebenen-Modell mit Random-Intercepts mit Random-Slopes von X1, X2 mit cross-level Effects
WITHIN ARE X1 X2; (1)	WITHIN ARE X1 X2; (1)	WITHIN ARE X1 X2; (1)
		BETWEEN ARE Z; (7)
CLUSTER=C2;	CLUSTER=C2;	CLUSTER=C2;
ANALYSIS: TYPE=TWOLEVEL;	ANALYSIS: (4) TYPE=TWOLEVEL RANDOM;	ANALYSIS: (4) TYPE=TWOLEVEL RANDOM;
MODEL: (2)  %WITHIN% Y1 ON X1 X2;	MODEL: (2)  %WITHIN% (5) beta1j beta2j   Y ON X1 X2;	MODEL:  %WITHIN% (5) beta1j beta2j   Y ON X1 X2;  %BETWEEN% Y beta1j beta2j ON Z; (8)
-3	-6	-9

(x) Die Anmerkungszahlen in der Tabelle dürfen nicht in die Mplus-Inputfiles übernommen werden.

(1) Hiermit werden X1 und X2 als L1-Prädiktoren definiert.

(2) Im 2-Ebenen-Modell ohne cross-level-Effekt(e) muss in Mplus keine BETWEEN-Anweisung angegeben werden.

(3) Die hier aufgeführten Mplus-Befehle werden in Modell 03 benutzt.

(4) Die Option "RANDOM" ist anzugeben, wenn Random-Slopes geschätzt werden sollen (so wie hier).

(5) Spezifikation von zwei Random-Slopes. Die ausführliche Version des Befehls wäre:

beta1j | Y ON X1;  
beta2j | Y ON X2;

(6) Die hier aufgeführten Mplus-Befehle werden in Modell 07 benutzt.

(7) Hiermit wird Z1 als L2-Prädiktor definiert.

(8) 3 Regressionen von Y, beta1j und beta2j auf die unabh. Var. "Z". Die ausführliche Version des Befehls wäre:  
Y ON Z;

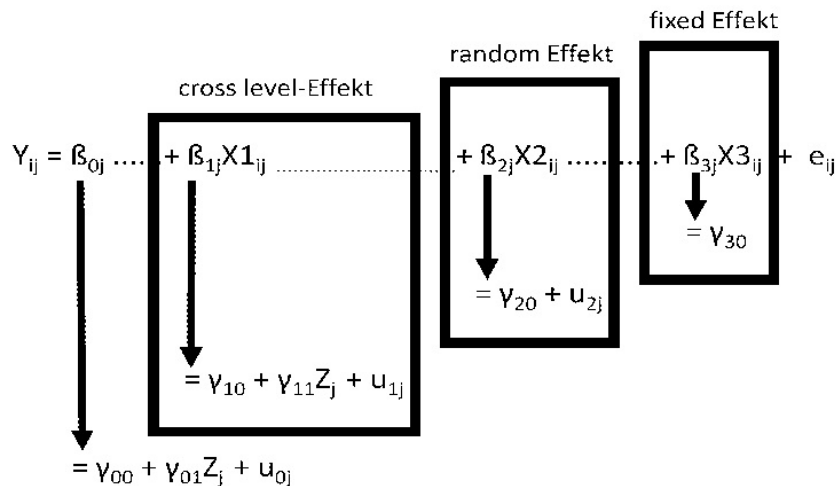
beta1j ON Z;  
beta2j ON Z;

(9) Die hier aufgeführten Mplus-Befehle werden in Modell 12 benutzt.

Die folgende Abbildung C2.2 macht noch einmal die drei unterschiedlichen Effekte des zuvor beschriebenen ML-Modells C2.1 deutlich.<sup>26</sup> Die drei in der Abbildung enthaltenen Rechtecke umranden (von rechts nach links) die Spezifikation:

- eines “fixed Effekt”,
- eines “random Effekt”,
- eines “cross level Effekt”.

Abb. C2.2: Drei Effekt-Typen im Zwei-Ebenen-Modell mit 3 X-Prädiktoren und 1 Z-Prädiktor



### C3 Standardisierung

Standardisierte Parameterschätzwerte sind problematisch.<sup>27</sup> Sie werden deshalb in diesem Skript nicht berechnet und auch nicht interpretiert.

Auch Mplus berichtet mit seinen Voreinstellungen keine standardisierten Parameterschätzwerte. Standardisierte Werte können aber in Mplus für einige MLA-Modellspezifikationen (nicht für alle!) mit einem Output-Befehl angefordert werden. Der dazu am häufigsten benutzte Output-Befehl lautet (hier zu verwenden u.a. in den Modellen 03 und 04):

STANDARDIZED(STDYX);

Es gibt drei verschiedene Möglichkeiten, um standardisierte Schätzwerte zu erhalten:

- 1) mit dem o.g. Mplus-Befehl
- 2) durch Transformation von unstandardisierten Schätzwerten in standardisierte Schätzwerte nach der folgenden Formel:<sup>28</sup>

<sup>26</sup> Zur besseren Veranschaulichung wurden die Terme in Abb. C2.2 auseinander gezogen und mit einer gepunkteten Linie untereinander verbunden.

<sup>27</sup> Vgl. dazu Urban/Mayerl 2018: 96 - 101.

<sup>28</sup> In der Formel ist “b” der unstandardisiert geschätzte Regressionskoeffizient (Slope) des X1-Effekts. Die Werte SD(x) und SD(y) sind die empirisch beobachteten Standardab-

$$b_{\text{StdYX}} = b \times [\text{SD}(x) / \text{SD}(y)]$$

Auch Hox et al. (2018: 18f) schlagen vor, für die Berechnung von standardisierten Effektkoeffizienten die unstandardisierten geschätzten Fixed-Parameter nachträglich nach der oben vorgestellten Formel zu standardisieren.

- 3) durch Standardisierung (auch “z-Transformation” genannt) der einzelnen empirischen Variablenwerte noch vor Durchführung der MLA nach der Formel:

$$Z_i = (X_i - \mu) / \sigma \quad (\text{mit: } X_i = \text{Messwert von } X_i, \mu = \text{Mittelwert von allen } X_i, \\ \sigma = \text{Standardabweichung von allen } X_i)$$

Diese Standardisierungsmöglichkeit sollte vor allem dann eingesetzt werden, wenn standardisierte cross-level-Effekte zu berechnen sind und Strategie (1) von Mplus nicht angeboten wird.

Die nach Standardisierung geschätzten Varianzkomponenten und Devianzen unterscheiden sich deutlich von den Werten, die ohne Standardisierung ausgegeben werden (denn sie sind nicht unabhängig von linearen Transformationen). Deshalb sollten bei MLA-Interpretationen die Schätzwerte für Varianzen und Devianzen (bzw. die Schätzwerte für den “random part” der Schätzung) auch immer von der unstandardisierten und nicht von der standardisierten Schätzung übernommen werden.

Da für die Standardisierung immer die empirischen Mittelwerte und Standardabweichungen der Modellvariablen benutzt werden, sollte darauf geachtet werden, dass zur Berechnung nur diejenigen Fälle benutzt werden, die auch in der multivariaten MLA eingesetzt werden. Wenn also für X1 einhundert gültige Messwerte vorliegen und für X2 nur neunzig Fälle mit gültigen Messwerten ermittelt wurden, dann muss die Standardisierung von X1 mit maximal neunzig Fällen (und nicht mit 100 Fällen) durchgeführt werden. In SPSS kann die Ermittlung gültiger Fälle u.a. mit den folgenden zwei Anweisungen erreicht werden:<sup>29</sup>

```
COUNT miss1 = x1 x2 ... (MISSING).
SELECT IF (miss1=0).
```

**VORSICHT:** Wie für alle standardisierten Schätzwerte gilt auch hier: Da, wie oben gezeigt, standardisierte Werte durch Verwendung von empirischen Mittelwerten und Standardabweichungen ermittelt werden, können sie nur für Vergleiche im selben Sample benutzt werden!

---

weichungen der unabh. Variablen “X1” und der abh. Variablen “Y”. Bei der Berechnung von standardisierten Werten ist darauf zu achten, dass die SD auf der richtigen Modellebene (Mikro- oder Makroebene) ermittelt werden.

<sup>29</sup> Ein anwendungsbezogenes Beispiel zur Ermittlung und Diskussion der multivariaten Missingquote wird in Urban/Mayerl 2018: 442-445 gegeben.

## **D Mehrebenen-Regressionsanalyse für 2 Ebenen**

### **D1 Die Gliederung**

Im Folgenden werden wir zunächst 14 Zwei-Ebenen-Regressionsmodelle vorstellen (die 9 Drei-Ebenen-Modelle werden dann in Kapitel E vorgestellt). Wir gehen davon aus, dass diese Modelle und die daraus abgeleiteten Modellspezifikationen die am häufigsten eingesetzten Modelle in der sozialwissenschaftlichen Zwei-Ebenen-Regressionsanalytik sind. Alle Modelle werden hier ausführlich erläutert, sodass ihre Logik, ihre Schätzung und ihre Interpretation im Kontext von Mplus-Input/Output-Files leicht verständlich werden.

Jedes Modell wird im Folgenden in der gleichen Art und Weise erläutert, sodass die Modelle leicht miteinander verglichen werden können. Zudem erfolgt zu Vergleichszwecken noch eine tabellarische Klassifikation aller 14 Zwei-Ebenen-Modelle (Kapitel D3).

Zu jedem der 14 Modelle werden Informationen in der folgenden Reihenfolge vermittelt:

1. Benennung des Modells mit seinen wichtigsten Kennzeichen (die dafür benutzten Textzeilen sind unterstrichen);
2. Auflistung aller formal-statistischen Gleichungen, die das Modell kennzeichnen;
3. Graphische Skizze des Gleichungssystems (Abbildung x.1);
4. Stilisierte graphische Darstellung der wichtigsten Kennzeichen des Modells und seiner zentralen Modell-Logik (Abbildung x.2);
5. Erläuterung der Modell-Logik und der Abbildung x.2 (vgl. Pkt. 4);
6. Interpretation des Gleichungssystems mit seinen einzelnen Modellgleichungen (vgl. Pkt. 2);
7. Hinweise zur Interpretation einzelner Ergebnisse der Modellschätzung;
8. Graphische Skizze des Gleichungssystems (vgl. Pkt. 2) mit den wichtigsten Schätzergebnissen (Abbildung x.3);
9. Tabellen mit den Ergebnissen der MLA-Schätzung (fixed und random coefficients);
10. Vollständige Syntax des Mplus-Inputfile mit integrierten Kommentaren;
11. Wichtige Ausschnitte aus dem Mplus-Outputfile mit integrierten Kommentaren.

Die für die einzelnen Komponenten der MLA benutzte Schreibweise (Notation) wird in Kapitel G näher vorgestellt. Allerdings dürften unsere Ausführungen auch ohne intime Kenntnis der formal-statistisch korrekten Notation verständlich sein.

Alle 14 Modellschätzungen benutzen den gleichen Datensatz. Diese Datensatz und die darin enthaltenen, analyserelevanten Variablen werden im folgenden Gliederungspunkt (D2) erläutert.

**D2 Der Datensatz**

Zur Schätzung der Zwei-Ebenen-Regressionsmodelle 01 bis 14 wird der Datensatz  
“kultur.DAT”

aus Luhmann (2015) benutzt. Dieser Datensatz ist kostenfrei von der Homepage:

“[www.beltz.de/index.php?id=1731](http://www.beltz.de/index.php?id=1731) (dort: 28790 Luhmann A5 Daten.zip)  
herunter zu laden.

Die Daten betreffen die Antworten von N=6929 befragten College-Studierenden (L1) in 40 Staaten/Ländern (L2).

Die durchschnittliche L2-Clustergröße (Cluster=Staat/Land) beträgt 173 befragte Personen.

Inhaltlich betrachtet wurde danach gefragt, ob die individuelle Lebenszufriedenheit der befragten Studierenden eher von deren sozio-psychischen Eigenheiten oder von deren wirtschaftlichen Lebensbedingungen bestimmt wird. Als Indikatoren werden das “Geschlecht” sowie die “Disposition zum Erleben positiver Gefühle” einer jeden Person benutzt. Als Indikator der ökonomischen Lebenssituation dient das “Bruttoinlandsprodukt” der 40 Staaten, in denen die befragten Studierenden leben.

Die hier benutzten Variablen des Datensatzes “kultur.DAT” sind:

(die Variablen: “pa2, sex2, gdp2“ werden in den jeweiligen Mplus-Inputfiles neu erstellt)

abh. Variable:	Y: lezu	Lebenszufriedenheit Skala: 1 ... 7
unabh. Variablen:	X1: pa	positiver Affekt = Häufigkeit des Erlebens positiver Gefühle Skala: 1 ... 9
	X1: pa2	Mediansplit von pa pa ≤ 5,67 → pa2=0 pa > 5,67 → pa2=1 (vgl. Mplus Input File zu Modell 08)
	X2: sex	1=male, 2=female
	X2: sex2	0/1=male/female (gebildet aus “sex” 1/2) (vgl. Mplus Input File zu Modell 03)
	Z: gdp	Bruttoinlandsprodukt (gross domestic product) Skala: 0,36 (Min.) ... 0,72 (Mean) ... 0,94 (Max)
	Z: gdp2	Bruttoinlandsprodukt nach Mediansplit 0 = gdp bis 0,74 1 = gdp über 0,74 (vgl. Mplus Input File zu Modell 05)
Cluster-ID:	C1: L1_ID	ID der jeweiligen Person bzw. der/des Studierenden
	C2: nation	ID des jeweiligen Landes/Staates



**D3 Tabellarische Klassifikation von 14 Zwei-Ebenen-Modellen**

<b>Zwei-Ebenen-Modelle:</b>		
	<b>Modelltyp:</b>	<b>Forschungsfrage(n):</b>
Modell 01	Random-Intercepts-Only-Modell -> ohne L1-Prädiktoren -> ohne L2-Prädiktoren	Wie groß ist der Anteil der L2-Varianz an der Gesamtvarianz von LEZU (die "Intra-Klassen-Korrelation", der "ICC-Koeffizient")? Ist der Anteil größer als 10% (Daumenregel), sodass eine MLA sinnvoll ist? (vgl. auch Kap. A1)
Modell 02	Random-Intercepts- /Fixed-Slopes-Modell -> mit L1-Prädiktor ("PA") -> ohne L2-Prädiktoren	Der Effekt des L1-Prädiktors "PA" (zentriert) auf die L1-Y-Variable "LEZU" (metrisch) soll unter länderspezifisch variierenden LEZU-Intercepts (=random intercepts) kontext-invariant (=fixed slopes) geschätzt werden.
Modell 03	Random-Intercepts- /Fixed Slopes-Modell -> mit L1-Prädiktor ("PA") -> mit L1-Prädiktor ("SEX2") -> ohne L2-Prädiktoren	Die länderspezifisch invarianten, partiellen Effekte (fixed) der L1-Prädiktoren "PA" (zentr.) und "SEX2" (dichotom) auf die L1-Y-Variable "LEZU" (metrisch) sollen geschätzt werden.
Modell 04	Random-Intercepts- /Random Slopes-Modell (intercepts-as-outcome model) -> ohne L1-Prädiktoren -> mit L2-Prädiktor ("GDP")	Der Effekt des L2-Prädiktors "GDP" (metrisch) auf die länderspezifisch variierenden LEZU-Intercepts (=random intercepts) soll kontext-variant (=random slopes) geschätzt werden.
Modell 05	Random-Intercepts- /Random-Slopes-Modell (intercepts-as-outcome model) -> ohne L1-Prädiktoren -> mit L2-Prädiktor ("GDP2")	... wie Modell 04, aber jetzt mit dichotomisiertem L2-Prädiktor "GDP2" (Mediansplit)
Modell 06	Random-Intercepts-/Random-Slopes-Modell -> mit L1-Prädiktor "PA" -> ohne L2-Prädiktoren	Der länderübergreifende Effekt des L1-Prädiktors "PA" (zentr.) auf L1-LEZU soll unter länderspezifisch variierenden LEZU-Intercepts und länderspezifisch variierenden PA-Effekten geschätzt werden
Modell 07	Random-Intercepts-/ Random-Slopes Modell -> mit L1-Prädiktor "PA" -> mit L1-Prädiktor "SEX2" -> ohne L2-Prädiktoren	Die partiellen Effekte des L1-Prädiktors "PA" (zentr.) und des L1-Prädiktors "SEX2" (dichotom) auf L1-LEZU sollen unter länderspezifisch variierenden PA- und SEX2-Effekten (beide:random) sowie länderspezifisch variierenden Intercepts (random) geschätzt werden und einen länderübergreifenden Schätzwert erhalten
Modell 08	Random-Intercepts-/ Random-Slopes Modell -> mit L1-Prädiktor "PA2" -> mit L1-Prädiktor "SEX2" -> ohne L2-Prädiktoren	... wie Modell 07, aber jetzt mit dichotomisiertem L1-Prädiktor "PA2" (Mediansplit)

Modell 09	Random-Intercepts- / Fixed + Random Slopes Modell -> mit L1-Prädiktor "PA2" -> mit L1-Prädiktor "SEX2" -> ohne L2-Prädiktoren	... wie Modell 08, aber jetzt mit Fixed-Slopes für SEX2-Effekt
Modell 10	Random-Intercepts- / Random Slopes Modell (intercepts-as-outcome model) -> mit L1-Prädiktor "PA" -> mit L2-Prädiktor "GDP2"	Es sollen länderübergreifende Schätzwerte ermittelt werden: a) für den L2-Effekt (random) von GDP2 auf die länderspezifischen LEZU-Intercepts, b) für den L1-Effekt (random) von PA auf LEZU.
Modell 11	Random-Intercepts- / Random Slopes Modell (intercepts-and-slopes-as-outcome model) -> mit L1-Prädiktor "PA" -> mit L2-Prädiktor "GDP2" -> mit "cross-level-effect"	Es sollen länderübergreifende Schätzwerte ermittelt werden: a) für den L2-Effekt (random) von GDP2 auf die länderspezifischen LEZU-Intercepts, b) für den L2-Effekt (random) von GDP2 auf den Slope (random) des L1-Einflusses von PA auf LEZU (cross-level-effect).
Modell 12	Random-Intercepts- / Random Slopes Modell (intercepts-and-slopes-as-outcome model) -> mit L1-Prädiktor "PA2" -> mit L1-Prädiktor "SEX2" -> mit L2-Prädiktor "GDP2" -> mit "cross-level-effects"	Es sollen länderübergreifende Schätzwerte ermittelt werden: a) für den Effekt (random) von GDP2 auf die länderspezifischen LEZU-Intercepts, b) für den L2-Effekt (random) von GDP2 auf den partiellen L1-Einfluss/Slope (random) von PA2 auf LEZU, c) für den L2-Effekt (random) von GDP2 auf den partiellen L1-Einfluss/Slope (random) von SEX2 auf LEZU.
Modell 13	Random-Intercepts- / Random Slopes Modell (intercepts-and-slopes-as-outcome model) -> mit L1-Prädiktor "PA2" -> mit L1-Prädiktor "SEX2" -> mit L2-Prädiktor "GDP2" -> mit "cross-level-effects"	.. wie Modell 12, aber jetzt ohne Effekt von GDP2 auf auf den partiellen Effekt/Slope (random) von SEX2 auf LEZU. Stattdessen erfolgt der partielle L1-Effekt (random) von SEX2 auf LEZU in direkter, ungebrochener Weise
Modell 14	Random-Intercepts- / Random Slopes Modell (intercepts-and-slopes-as-outcome model) -> mit L1-Prädiktor "PA2" -> mit L1-Prädiktor "SEX2" -> mit L2-Prädiktor "GDP2" -> mit "cross-level-effects"	... wie Modell 13, aber mit fixed SEX2-Effekten

#### **D4 Systematische Erläuterung von 14 Zwei-Ebenen-Modellen** (Modell 01 bis Modell 14)

Modell 01: Nullmodell, ohne Prädiktoren (random-intercepts-only modell)	48
Modell 02: ein L1-Prädiktor-Effekt (fixed) auf Intercepts (random) (random intercepts/fixed slopes modell)	55
Modell 03: wie Modell 02, aber mit zwei fixen L1-Prädiktor-Effekten auf Intercepts (random) (random intercepts/fixed slopes model)	61
Modell 04: ein L2-Prädiktor-Effekt (random) auf Intercepts (random) (random intercepts as outcome/random slopes modell)	69
Modell 05: wie Modell 04, aber mit dichotomen L2-Prädiktor (random). (random intercepts as outcome/random slopes modell)	75
Modell 06: wie Modell 02, aber mit einem random L1-Prädiktor-Effekt. (random intercepts/random slopes modell)	81
Modell 07: wie Modell 06, aber mit zwei partiellen L1-Prädiktor-Effekten. (random intercepts/random slopes modell)	87
Modell 08: wie Modell 07, aber jetzt sind beide L1-Prädiktoren dichotom (random intercepts/random slopes modell)	94
Modell 09: wie Modell 08, aber jetzt mit einem fixed und einem random L1-Prädiktor (random intercepts/fixed+random slopes modell)	101
Modell 10: Modell mit einem L1-Prädiktor (random) und einem L2-Prädiktor (random) (random intercepts as outcome/random slopes modell)	109
Modell 11: ein L2-Prädiktor-Effekt (random) auf Intercepts und ein Effekt (random) eines L2-Prädiktors auf einen L1-Slope (cross-level-effect, random intercepts and random slopes as outcome)	115
Modell 12: ein L2-Prädiktor-Effekt (random) auf Intercepts und auf zwei L1-Slopes (cross-level-effects, random intercepts and random slopes as outcome)	122
Modell 13: wie Modell 12, aber nur mit einem cross-level-effect (random) (cross-level-effects, random intercepts and random slopes as outcome)	132
Modell 14: wie Modell 13, aber mit einem L1-Prädiktor-Effekt (fixed) (cross-level-effect, random intercepts and random slopes as outcome)	141

Modell 01:

Zwei-Ebenen-Modell,

Random-Intercepts-Only-Modell,

ohne L1/L2-Prädiktoren

L1-Gleichung:  $Y_{ij} = \beta_{0j} + e_{ij}$  (1.1)

L2-Gleichung:  $\beta_{0j} = \gamma_{00} + u_{0j}$  (1.2a)

“fixed part”:  $\gamma_{00}$

“random part”:  $\text{Var}(e_{ij}), \text{Var}(u_{0j})$

(Das Modell 01 entspricht dem Modell 3.0 und dessen Formalisierung in: Finch/Bolin 2017: 47 - 52).

Abb. 1.1: Graphische Skizze des Gleichungssystems von Modell 01

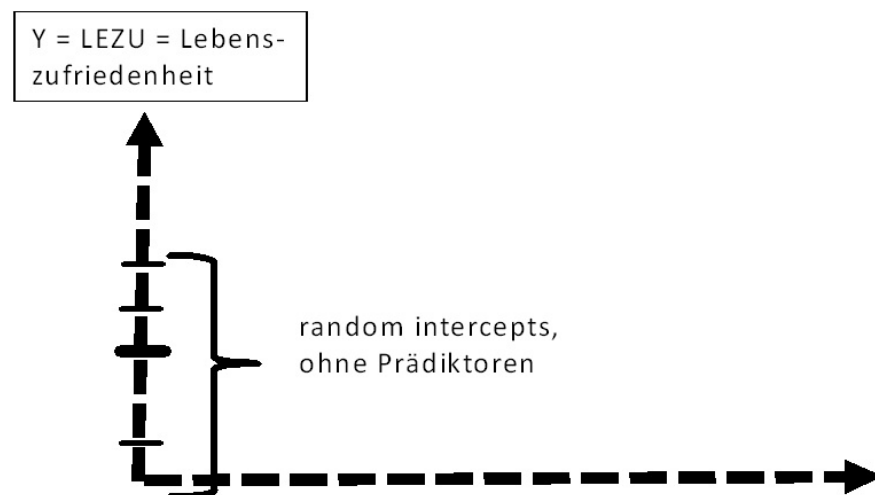
$$Y_{ij} = \beta_{0j} + e_{ij}$$

↓

$$= \gamma_{00} + u_{0j}$$

$\beta$ : beta     $\gamma$ : gamma  
Y: LEZU

Abb. 1.2: Stilisierte Darstellung der Modell-Logik von Modell 01



Zur Logik von Modell 01:

Das Random-Intercepts-Only-Modell (auch: Null-Modell, Baseline-Modell) hat keine Prädiktoren (es werden allein durchschnittliche Y-Werte geschätzt). Das Modell dient zur Aufteilung der Gesamtvarianz von kontextuell eingebetteten Y-Messungen in eine L1-Varianz (“within” Varianz auf der Individualebene) und eine L2-Varianz (“between” Varianz auf der Kollektivebene). Mit diesen Teilvarianzen lässt sich die (problematische) Intra-Klassen-Korrelation (ICC-Koeffizient) berechnen, die angibt, wie groß der prozentuale Anteil der L2-Varianz an der Gesamtvarianz ist (vgl. Kap. A1). Mit der Modellschätzung soll ermittelt werden:

- a) wie hoch der ICC-Koeffizient ist,
- b) ob der Anteil der Between-Varianz an der Gesamtvarianz groß genug ist (>10%), um eine MLA mit L1- und L2-Prädiktoren durchzuführen.

Die Abbildung 1.2 verdeutlicht die Logik des Modells 01. Die einzelnen länderspezifischen Mittelwerte von  $\bar{Y}=\text{LEZU}$  liegen auf der vertikalen Y-Achse. Ihre Lage wird dort als kurzer Längsstrich markiert. Der breite kurze Längsstrich auf der Y-Achse markiert die Lage des länderspezifischen Gesamtmittelwertes (grand mean). Die länderspezifischen Mittelwerte sind identisch mit den länderspezifischen intercepts in der Gleichung 1.1. Die länderspezifischen intercepts sind “random”, weil die Länder unterschiedlich große Y-Werte aufweisen. Da im Modell keine Prädiktoren berücksichtigt werden, steht die vertikale Y-Achse ganz links auf der horizontalen X-Achse, wo X-Variablen bzw. X-Prädiktoren einen Wert von 0 hätten. Die Abstände zwischen den länderspezifischen Y-Mittelwerten und dem länderspezifischen Gesamtmittelwert symbolisieren die Größe der länderspezifischen Residualvariablen  $u_{0j}$  in Gleichung 1.2a.

(Ausführlichere Informationen zur Logik von MLA-Modellen finden sich bei der Erörterung von Modell 20).

Zur Interpretation des Gleichungssystems von Modell 01:

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (1.1)$$

Y Die Variable Y ist die abh. Variable auf der Mikroebene L1. In unserem Datensatz ist  $Y=\text{LEZU}$ .

$\beta_{0j}$  Für jeden Befragten wird ein LEZU-Wert geschätzt, der identisch mit dem LEZU-Mittelwert aller Befragten eines jeweiligen Staates/Landes ist. Dieser Mittelwert ist identisch mit dem Achsenabschnitt  $\beta_{0j}$ .

$e_{ij}$  Der mit  $\beta_{0j}$  geschätzte bzw. der gemittelte Y-Wert soll für jede/n Befragte/n innerhalb eines bestimmten Landes gelten. Der tatsächliche Wert einer jeden Person entspricht aber oftmals nicht ihrem tatsächlichen Y-Wert. Der jeweilige reale Wert wird oberhalb oder unterhalb des geschätzten Achsenabschnitts liegen. Die Differenz zwischen geschätztem und empirischen Y-Wert wird als “e” bezeichnet. Diese Abweichungen/Differenzen werden auch als L1-Residuen bezeichnet.

$\text{Var}(e_{ij})$  Ist die Varianz der Residualvariable  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster/Staaten/Länder).

Mplus: Within Level / Variances LEZU = 1,229 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (1.2a)$$

$\gamma_{00}$  Die Gleichung 1.2a enthält den Gesamtmittelwert “ $\gamma_{00}$ ” (auch: Grand Mean) aller einzelnen, länderspezifischen Mittelwerte.

Mplus: Between Levels / Means LEZU = 4,396 (p=0,000)

$u_{0j}$  Ist die Residualvariable auf der L2-Ebene. Die Variablenwerte entsprechen den Abweichungen der länderspezifischen Mittelwerte/Achsenabschnitte “ $\beta_{0j}$ ” vom Gesamtmittelwert “ $\gamma_{00}$ ”.

$\text{Var}(u_{0j})$  Ist die Varianz der Residualvariable  $u_{0j}$  (s.o.) auf der L2-Ebene. Wird auch als Between-Varianz bezeichnet (Varianz zwischen den Clustern/Staaten/Ländern).

Mplus: Between Level / Variances LEZU = 0,279 (p=0,000)

### Zur Interpretation der Schätzung von Modell 01:

(vgl. dazu auch die folgende Abbildung 1.3 sowie die folgenden Tabellen 1.1 und 1.2)

- 1) Die beiden Residualvariablen  $e_{ij}$  (L1-Ebene) und  $u_{0j}$  (L2-Ebene) können als zufällige Größen (random coefficients) bezeichnet werden, da sie mit Auswahl eines jeweiligen Befragten “i” und/oder eines jeweiligen Landes “j” variieren (Eselsbrücke: es gibt für beide Variablen auch Subskripte i und/oder j).
- 2) Mit den Varianzen aus dem Random-Intercept-Only-Modell kann auch die Intra-Klassen-Korrelation (ICC) für das jeweilige Mehrebenen-Modell berechnet werden (vgl. auch Kap. A1):

$$\text{Var}(u) / [\text{Var}(u) + \text{Var}(e)] \text{ oder: } \text{L2-Varianz} / [\text{L2-Varianz} + \text{L1-Varianz}].$$

Allerdings wird der ICC-Koeffizient vom Mplus-Output auch direkt ausgegeben (s.u.). Der ICC informiert über den Anteil der L2-Varianz “ $\text{Var}(u)$ ” an der Gesamtvarianz und sollte für eine MLA mindestens einen Wert von 0,10 (10%) aufweisen, d.h. der Anteil der L2-Varianz an der Gesamtvarianz sollte größer sein als 10%, damit es sinnvoll ist, eine MLA durchzuführen. In unserem Beispiel beträgt der ICC-Koeffizient “0,185”. Somit ist die Clusteraufteilung der Daten für 18,5% der Gesamtvarianz in unserem Modell verantwortlich. Es sollte mithin eine MLA durchgeführt werden, um L2-Prädiktoren zu identifizieren, die weitere Anteile der L2-Varianz binden können.

Mplus: Variable LEZU / Intraclass Correlation = 0,185

- 3) Die L1-Varianz und die L2-Varianz sind hochgradig signifikant (vgl. Tab. 1.2). Es können also durchaus weitere L1/L2-Prädiktoren getestet werden, um freie Varianzanteile zu binden. Die signifikante L2-Varianz verweist darauf, dass es beträchtliche Abweichungen der länderspezifischen LEZU-Mittelwerte vom Gesamtmittelwert (Grand Mean) gibt.

Abb. 1.3: Graphische Skizze des Gleichungssystems von Modell 01 mit Schätzergebnaten.

$$\begin{aligned} \text{LEZU}_{ij} &= \beta_{0j} + e_{ij} \\ &\downarrow \\ &= \gamma_{00} + u_{0j} \\ &\quad (4,39^{***}) \end{aligned}$$

Tab. 1.1: Ergebnisse der MLA-Schätzung von Modell 1: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept ( $\gamma_{00}$ ) Mean(LEZU)	4,396	0,000

Tab. 1.2: Ergebnisse der MLA-Schätzung von Modell 1: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Varianz, LEZU, Var( $e_{ij}$ )	1,229	0,000
L2-Varianz: LEZU, Intercept, Var( $u_{0j}$ )	0,279	0,000

Tab. 1.3: Mplus-Inputfile für Modell 01

```
1 TITLE:
2   Modell "MLA 01"
3   Baseline Model, Null-Modell
4   keine Prädiktoren (nur Intercepts)
5
6 DATA:
7   File is "E:/data/mplus/kultur.dat" ;
8
9 VARIABLE:
10  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
11            KLA_F  LEZU  PA  NA  BAL  IC
12            GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
13            PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;
14
15 MISSING ARE . ;
16
17 USEVARIABLES ARE
18   nation ! nation ist die L2-Cluster-Variable
19   lezu; ! lezu ist die abh. Y-Variable
20         ! die Reihenfolge der Variablen ist bedeutungslos
21
22 CLUSTER=nation;
23
24 ANALYSIS:
25   TYPE = TWOLEVEL;
26
27 ! MODEL: ist nicht notwendig, da es hier keine L1/L2-Prädiktoren gibt
28
```



Tab. 1.4: Mplus-Ausgabe von Modellschätzung 01 (ausgewählte Bereiche)

1	SUMMARY OF DATA			
2	Number of missing data patterns		1	
3	Number of clusters	40		
4	Average cluster size	173.225		
5				
6	Estimated Intraclass Correlations for the Y Variables			
7	Intraclass			
8	Variable	Correlation		
9	LEZU	0.185		
10				
11	MODEL FIT INFORMATION			
12	Number of Free Parameters		3	
13				
14	Loglikelihood			
15	H0 Value	-10616.565		
16	H0 Scaling Correction Factor for MLR	1.8492		
17	H1 Value	-10616.565		
18	H1 Scaling Correction Factor for MLR	1.8492		
19	Information Criteria			
20	Akaike (AIC)	21239.129		
21	Bayesian (BIC)	21259.660		
22	Sample-Size Adjusted BIC	21250.126	(n* = (n + 2) / 24)	
23	Chi-Square Test of Model Fit			
24	Value	0.000*		
25	Degrees of Freedom	0		
26	P-Value	1.0000		
27	Scaling Correction Factor for MLR	1.0000		
28				
29	* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used			
30	for chi-square difference testing in the regular way. MLM, MLR and WLSM			
31	chi-square difference testing is described on the Mplus website. MLMV, WLSMV,			
32	and ULSMV difference testing is done using the DIFFTEST option.			
33				
34	RMSEA (Root Mean Square Error Of Approximation)			
35	Estimate	0.000		
36	CFI/TLI			
37	CFI	0.000		
38	TLI	1.000		
39	Chi-Square Test of Model Fit for the Baseline Model			
40	Value	0.000		
41	Degrees of Freedom	0		
42	P-Value	1.0000		
43	SRMR (Standardized Root Mean Square Residual)			
44	Value for Within	0.000		
45	Value for Between	0.000		
46	MODEL RESULTS			
47		Two-Tailed		
48	Estimate	S.E. Est./S.E.	P-Value	
49				
50	Within Level			
51			FORTSETZUNG→	
52				

53

54 Fortsetzung von Tabelle 1.4:

55

56

57 Variances

58 LEZU 1.229 0.040 30.470 0.000

59

60 Between Level

61

62 Means

63 LEZU 4.396 0.085 51.810 0.000

64

65 Variances

66 LEZU 0.279 0.059 4.720 0.000

67



Zur Logik von Modell 02:

Im Random-Intercepts-/Fixed-Slopes-Modell (Modell 02) wird das Random-Intercepts-Only-Modell (Modell 01) um den Prädiktor X1 auf der L1-Ebene erweitert (also um einen Prädiktor dessen Ausprägungen zwischen den Befragten variiert). Mit der Modellschätzung soll ermittelt werden:

- wie stark der Effekt von X1 auf Y ist, wenn die Zwei-Ebenen-Struktur der Daten berücksichtigt wird und gleichzeitig angenommen wird, dass der Effekt in allen Ländern gleich ist (fixed effect),
- wie hoch der Durchschnittswert von Y ist.

Die Abbildung 2.2 verdeutlicht die Logik von Modell 02. Auf der Y/LEZU-Achse sind die unterschiedlichen länderspezifischen Intercepts der Gleichung 2.1 abgetragen. Diese Werte sind demnach "random" verteilt. Die durchgezogenen Linien im gestrichelt gezeichneten Koordinatensystem markieren die länderspezifischen Regressionsgeraden nach Gleichung 2.1 bzw. die länderspezifischen Effekte von X1/PA auf Y/LEZU. Die Steigungen dieser Geraden sind alle gleich groß. Deshalb sind sie "fixed slopes".

Zur Interpretation des Gleichungssystems von Modell 02:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + e_{ij} \quad (2.1)$$

Die Gleichung 2.1 ist eine L1-Gleichung zur Schätzung von Y=LEZU durch X1=PA.

Y abh. Variable LEZU auf der Mikroebene L1

$\beta_{0j}$  Achsenabschnitt der Schätzung von Y (s.u.)

$\beta_{1j}$  Slope des Effekts von X1 auf Y auf der Mikroebene (s.u. " $\gamma_{10}$ ")

X1 unabh. Variable PA auf der Mikroebene L1 (zentriert mit Grand Mean)

$e_{ij}$  Residualvariable, Diff. zwischen beob. LEZU und geschätzter LEZU auf Befragtenebene

Var( $e_{ij}$ ) Varianz der Residualvariable  $e_{ij}$ . Es wird angenommen, dass auf jeder Ebene (also hier auf L1 und L2) die Residuen (hier: e und u) einen Mittelwert von 0 und eine konstante Varianz für alle Werte von X haben.

Mplus: Within Level / Residual Variances LEZU: 1.007 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2.2a)$$

Die L2-Gleichung dient der Ermittlung von  $\beta_{0j}$ .

$\gamma_{00}$  Zentraler Achsenabschnitt (mittlerer Wert aller länderspezifischen Intercepts). Der zentrale Achsenabschnitt  $\gamma_{00}$  berichtet das zu erwartende Ausmaß von Y, wenn der Prädiktor X1 einen Wert von 0,00 hätte. Um diesen Wert von 0,00 besser verstehen zu können, werden in der MLA üblicherweise die unabh. X-Prädiktoren noch vor der ML-Analyse zentriert (vgl. Kap. B4), sodass X1=0.00 dem empirischen Durchschnittswert von X entspricht (s.o. "X1").

Mplus: Between Level / Means LEZU: 4,423 (p=0.000).

$u_{0j}$  L2-Residualvariable: die länderspezifischen Achsenabschnitte variieren um den zentralen Achsenabschnitt von  $\gamma_{00}$ .

Var( $u_{0j}$ ) Varianz der L2-Residualvariable  $u_{0j}$  (s.o.)

Mplus: Between Level / Variances LEZU: 0,186 (p=0.000)

$$\beta_{1j} = \gamma_{10} \quad (2.2b)$$

$\beta_{1j}$  Schätzung eines über die Länder hinweg konstanten PA-Slopes (= fixed slope, s.u.).

$\gamma_{10}$  Die Effektstärke, mit der X1 (=PA) die abh. Variable Y (=LEZU) beeinflusst ( $\beta_1$ ), ist im Prinzip länderspezifisch ( $\beta_{1j}$ ) und kann wie ein unstandardisierter Regressionskoeffizient interpretiert werden. Allerdings gehen wir für das Modell 2 davon aus, dass die Effektstärke nicht variiert, sondern in jedem Cluster eher gleich ist und der festen L2-Größe " $\gamma_{10}$ " entspricht. Die L2-Größe " $\gamma_{10}$ " ist somit, so die Annahme, relativ homogen über alle L2-Einheiten (fixed slope-Annahme).

Mplus: Within Level / LEZU ON PA: 0,351 (p=0,000)

#### Zur Interpretation der Schätzung von Modell 02:

(vgl. dazu auch die folgende Abbildung 2.3 sowie die Tabellen 2.1 und 2.2)

- 1) Alle Schätzergebnisse beruhen auf der Annahme, dass der Effekt von PA auf LEZU in allen Ländern gleich ist. Dies ist aber sehr unwahrscheinlich. Die Annahme wird hier nur getroffen, um systematisch in die MLA einzuführen.
- 2) Die Stärke des Effekts von PA auf LEZU beträgt 0,351 und ist hochsignifikant. Wenn sich PA um eine empirische Einheit erhöht, steigt der zu erwartende LEZU-Wert um 0,35 Punkte (LEZU-Skala: 1,00-7,00).
- 3) Der durchschnittliche LEZU-Wert beträgt dann 4,42, wenn PA gleich null ist. Da PA zentriert wurde, ist PA=0, wenn Befragte den Durchschnittswert von PA aufweisen.
- 4) Die Varianz des L2-Intercepts ist beträchtlich und hochsignifikant. Das Intercept variiert deutlich von Land zu Land. Ähnliches kann man von den PA-Slopes nicht sagen, denn diese sind ja auf einen festen Wert fixiert. Allerdings signalisieren die AIC/BIC-Werte von Modell 02, die deutlich kleiner sind als diejenigen von Modell 01, dass schon eine Erweiterung des Modells um einen L1-Prädiktor zu einer wesentlichen Verbesserung der Anpassung des Modells führt, auch wenn die länderspezifischen PA-Slopes fixiert sind (vgl. Kap. B5).

Abb. 2.3: Graphische Skizze des Gleichungssystems von Modell 02 mit Schätzergebnaten

$$Y_{ij} = \beta_{0j} + \dots + \beta_{1j}(X1_{ij}) + e_{ij}$$

$\beta_{0j}$  ↓  $\beta_{1j}$   
 $= \gamma_{00} + u_{0j}$   $= \gamma_{10}$   
 (4,42\*\*\*) (0,35\*\*\*)

$\beta$ : beta  
 $\gamma$ : gamma  
 $Y$ : LEZU  
 $X1$ : PA\_zentr

Tab. 2.1: Ergebnisse der MLA-Schätzung von Modell 02: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept ( $\gamma_{00}$ ) Mean(LEZU)	4,423	0,000
L1-Slope (fixed), ( $\gamma_{10}$ ) LEZU ON PA_zentr	0,351	0,000

Tab. 2.2: Ergebnisse der MLA-Schätzung von Modell 02: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	1,007	0,000
L2-Varianz, LEZU, Intercept, Var( $u_{0j}$ )	0,186	0,000

Tab. 2.3: Mplus-Inputfile für Modell 02

```
TITLE:
  Modell 02 (verbessertes Input-File)
  mit random intercepts, mit einem L1-Prädiktor "pa"
  mit fixed slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu ! lezu ist die abh. Y-Variable
  pa; ! pa ist X-Prädiktor (wird später noch zentriert)

WITHIN = pa; ! hiermit wird pa als L1-Prädiktor definiert

BETWEEN =; ! kein L2-Prädiktor

CLUSTER = nation;

DEFINE:
  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert
                        ! (keine Zentrierung bei Dummies)

ANALYSIS:
  TYPE = TWOLEVEL;

MODEL:
  %BETWEEN%
    lezu;
  %WITHIN%
    lezu ON pa;

OUTPUT:
  sampstat;
```

Tab. 2.4: Mplus-Ausgabe von Modellschätzung 02 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters		4		
Loglikelihood				
H0 Value		-9921.914		
H0 Scaling Correction Factor for MLR		2.4491		
H1 Value		-9921.914		
H1 Scaling Correction Factor for MLR		2.4491		
Information Criteria				
Akaike (AIC)		19851.828		
Bayesian (BIC)		19879.202		
Sample-Size Adjusted BIC		19866.491		
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
LEZU	ON			
PA	0.351	0.018	19.407	0.000
Residual Variances				
LEZU	1.007	0.034	29.857	0.000
Between Level				
Means				
LEZU	4.423	0.070	63.248	0.000
Variances				
LEZU	0.186	0.040	4.604	0.000



Modell 03:

Random-Intercepts-/Fixed-Slopes-Modell,

mit L1-Prädiktor "PA" (zentriert),

mit L1-Prädiktor "SEX2" (dichotomisiert, 0/1 = m/w),

ohne L2-Prädiktoren

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (3.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + u_{0j}$  (3.2a)

$\beta_{1j} = \gamma_{10}$  (3.2b)

$\beta_{2j} = \gamma_{20}$  (3.2c)

"fixed part":  $\gamma_{00}, \gamma_{10}, \gamma_{20}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j})$

Abb. 3.1: Graphische Skizze des Gleichungssystems von Modell 03

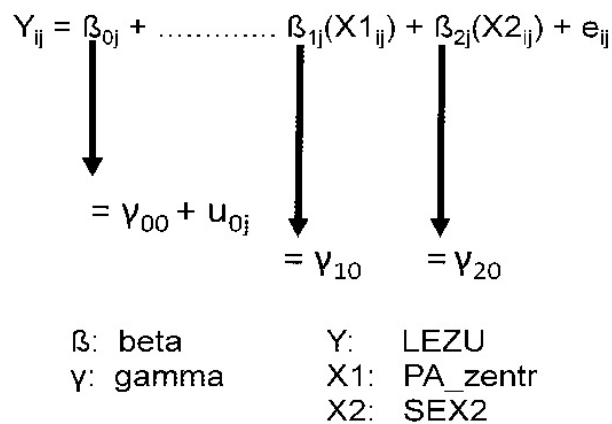


Abb. 3.2: Stilisierte Darstellung der Modell-Logik von Modell 03

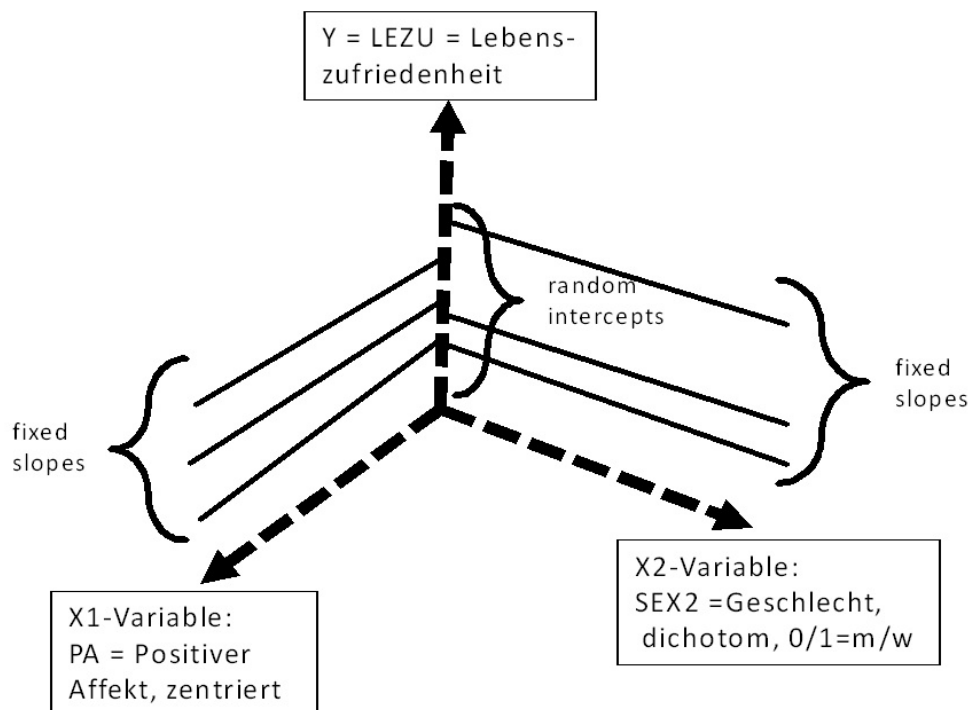
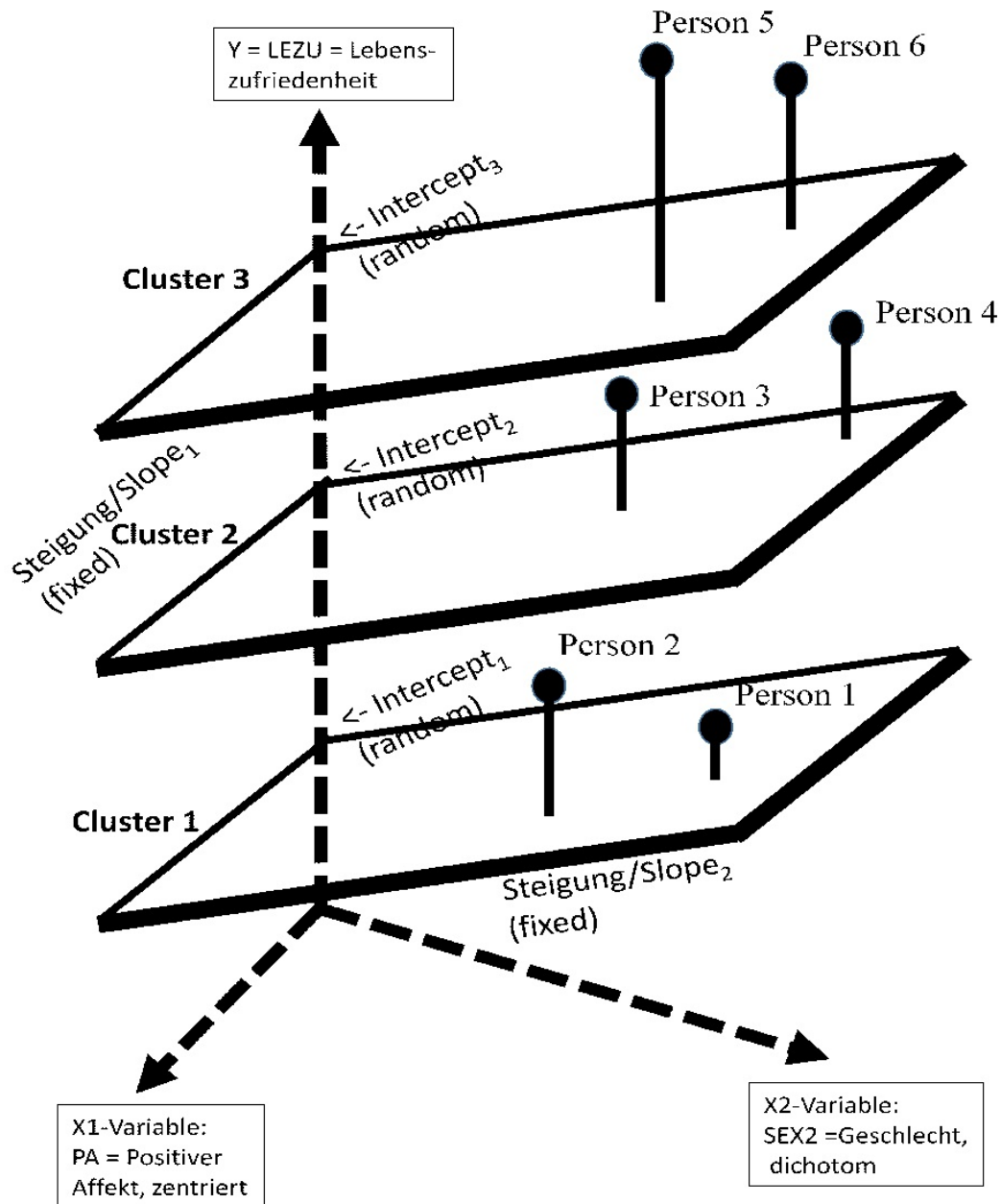


Abb. 3.3: Stilisierte Darstellung der Modell-Logik von Modell 03



#### Zur Logik von Modell 03:

Die Struktur von Modell 03 ist identisch mit derjenigen von Modell 02. Der große Unterschied zwischen beiden Modellen ist die Anzahl der L1-Prädiktoren. In Modell 03 gibt es neben dem  $X_1$ -Prädiktor "PA" auch noch den  $X_2$ -Prädiktor "SEX2 bzw. FEMALE" (dichotom, 0/1=m/w). Dementsprechend muss die L1-Gleichung 3.1 um den zusätzlichen Prädiktor SEX2 und dessen Regressionskoeffizient erweitert werden.

Die L1-Gleichung ist identisch mit einer multiplen Regressionsgleichung. Somit können die

beiden Einflusskoeffizienten auch als partielle Regressionskoeffizienten interpretiert werden. Zudem können Interaktionen zwischen den Prädiktoren modelliert werden. Viele Modellannahmen multipler OLS-Regressionsmodelle gelten auch für die L1-Gleichung 3.1 und eine diesbezügliche ML-Schätzung (z.B. Abwesenheit von Multikollinearität, vgl. Kap. A2, Pkt.7 ). Mit der Modellschätzung soll ermittelt werden:

- a) wie stark die Effekte von X1 und X2 auf Y sind, wenn die Zwei-Ebenen-Struktur der Daten berücksichtigt wird und gleichzeitig angenommen wird, dass die beiden Effekte in allen Ländern gleich sind (fixed effect),
- b) welcher Effektkoeffizient der stärkere ist,
- c) wie hoch der Durchschnittswert von Y ist.

Die Abbildungen 3.2 und 3.3 verdeutlichen die Logik von Modell 03. Beide Abbildungen verweisen auf den gleichen Sachverhalt. Im Folgenden konzentrieren wir uns auf Abbildung 3.3: Die Abbildung 3.3 zeigt ein dreidimensionales Koordinatensystem (gestrichelte Linien). Die drei Koordinaten beziehen sich auf die drei Variablen Y, X1, X2 bzw. LEZU, PA, SEX2. Im Koordinatensystem sind drei Flächen angeordnet. Diese Flächen beziehen sich auf die Cluster 1 bis 3 bzw. auf die Länder/Staaten 1 bis 3. Zusätzlich sind für jede Fläche bzw. für jedes Cluster/Land zwei Personen eingezeichnet (fette Punkte), deren jeweilige Lage im dreidimensionalen Koordinatensystem sich aus ihren empirischen Werten auf den Variablen Y, X1 und X2 ergibt. Für das Cluster/Land 2 sind das z.B. die Punkte, welche die empirische Lage von Person 5 und Person 6 anzeigen

Mit Gleichung 3.1 kann die erwartete Lage einer jeden Person im Koordinatensystem geschätzt werden. Die geschätzten Positionen liegen für jede Person auf den eingezeichneten Flächen und sind genau dort verortet, wo die senkrechte Linie, die jeden Punkt nach unten verlässt, auf die betreffende Fläche stößt. Der Abstand der Personenpunkte zur geschätzten Fläche verdeutlicht somit die Differenz zwischen beobachteten und geschätzten Personenwerten. Je länger die Linie ist, die vom Punkt zur Fläche reicht, umso größer sind die Differenzen zwischen beobachteten und geschätzten Personenwerten.

Jede Fläche stößt mit einer ihrer Ecken an die Y-Linie (gestrichelt) des Koordinatensystems. Diese Eckpunkte markieren die Intercept-Werte der Gleichung 3.1 für jedes einzelne Land. Der Eckpunkt für Cluster/Land 2 liegt oberhalb des Eckpunkts von Cluster/Land 1 und folglich ist das Intercept von Cluster/Land 2 größer als dasjenige von Cluster/Land 1. Am größten ist mithin das Intercept von Cluster/Land 3. Da die Intercepts länderspezifisch variieren werden sie auch als "random" Werte bezeichnet. Die Intercepts geben die geschätzten LEZU-Werte für den Fall an, dass die Y-Variablen einen Wert von null annehmen. Das kann z.B. dann auftreten, wenn eine Variable zentriert ist (wie in Modell 03 die Variable PA) und ihr Mittelwert einen Wert von null annimmt.

Jede Fläche weist zwei unterschiedliche und voneinander unabhängige Steigungen auf: einmal entlang der X1-Koordinate und einmal entlang der X2-Koordinate. Diese Steigungen sind positiv. D.h.: Je größer der geschätzte Wert für X1/PA bzw. für X2/SEX2 wird, umso größer wird der geschätzte Wert für Y/LEZU. Beide Steigungsraten gelten für alle untersuchten Länder (hier: für Cluster/Land 1 bis 3). Sie variieren also nicht zwischen den verschiedenen Ländern. Deshalb werden die Steigungsraten bzw. die Slopes auch als "fixed" bezeichnet.

#### Zur Interpretation des Gleichungssystems von Modell 03:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (\text{Gl. 3.1})$$

Die L1-Gleichung 3.1 dient der Schätzung von Y=LEZU.

Y abh. Variable LEZU auf der Mikroebene L1

$\beta_{0j}$  Achsenabschnitt der Schätzung von Y (s.u.)

$\beta_{1j}$  Slope des Effekts von X1 auf Y auf der Mikroebene (s.u. " $\gamma_{10}$ ")

X1 unabh. Variable PA auf der Mikroebene L1 (zentriert mit Grand Mean)

$\beta_{2j}$  Slope des Effekts von X2 auf Y auf der Mikroebene (s.u. " $\gamma_{10}$ ")

X2 unabh. Variable SEX2 (dichotomisiert, 0/1=m/w) auf der Mikroebene L1

$e_{ij}$  Residualvariable, Differenz zwischen beobachteter LEZU und geschätzter LEZU auf Befragtenebene

Var( $e_{ij}$ ) Varianz der Residualvariable  $e_{ij}$

Mplus: Within Level / Residual Variances LEZU: 1.003 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (\text{Gl. 3.2a})$$

Die L2-Gleichung 3.2a wird zur Ermittlung von  $\beta_{0j}$  eingesetzt.

$\gamma_{00}$  Zentraler Achsenabschnitt (mittlerer Wert aller länderspezifischen Intercepts). Der zentrale Achsenabschnitt  $\gamma_{00}$  berichtet das zu erwartende Ausmaß von Y, wenn die Prädiktoren X1 und X2 einen Wert von 0,00 haben. Um diesen Wert von 0,00 besser verstehen zu können, werden in der MLA üblicherweise die unabh. metrischen X-Prädiktoren noch vor der ML-Analyse zentriert, sodass X1=0.00 den empirischen Durchschnittswert von X1 markiert (s.o. "X1"). Der Prädiktor X2 hat einen Wert von 0 bei einem Geschlecht von männlich (s.o. "X2").

Mplus: Between Level / Means LEZU: 4,350 (p=0.000).

$u_{0j}$  L2-Residualvariable: die länderspezifischen Achsenabschnitte variieren um den zentralen Achsenabschnitt von  $\gamma_{00}$  und die Differenzen zwischen dem mittleren Achsenabschnitt und den länderspezifischen Achsenabschnitten werden hier als  $u_{0j}$  bezeichnet.

Var( $u_{0j}$ ) Varianz der L2-Residualvariable  $u_{0j}$  (s.o.)

Mplus: Between Level / Variances LEZU: 0,179 (p=0.000)

$$\beta_{1j} = \gamma_{10} \quad (3.2b)$$

Schätzung eines über die Länder hinweg konstanten PA-Slopes (= fixed slope).

$\gamma_{10}$  Die Effektstärke, mit der X1 (=PA) die abh. Variable Y (=LEZU) beeinflusst ( $\beta_1$ ), ist im Prinzip länderspezifisch ( $\beta_{1j}$ ) und kann wie ein unstandardisierter Regressionskoeffizient interpretiert werden. Allerdings wird mit Gleichung 3.2b festgelegt, dass die Effektstärke nicht variiert, sondern in jedem Land gleich ist und dabei der festen L2-Größe " $\gamma_{10}$ " entspricht (fixed slope-Annahme).

Mplus: Within Level / LEZU ON PA: 0,348 (p=0,000)

$$\beta_{2j} = \gamma_{20} \quad (3.2c)$$

Schätzung eines über die Länder hinweg konstanten PA-Slopes (= fixed slope).

$\gamma_{20}$  Die Effektstärke, mit der X2(=SEX2) die abh. Variable Y (=LEZU) beeinflusst ( $\beta_2$ ), ist im Prinzip länderspezifisch ( $\beta_{2j}$ ) und kann wie ein unstandardisierter Regressionskoeffizient interpretiert werden. Allerdings wird mit Gleichung 3.2c festgelegt, dass die Effektstärke nicht variieren soll, sondern in jedem Land gleich ist und dabei der festen L2-Größe " $\gamma_{20}$ " entspricht (fixed slope-Annahme).

Mplus: Within Level / LEZU ON PA: 0,122 (p=0,000)

### Zur Interpretation der Schätzung von Modell 03

(vgl. dazu auch die folgende Abbildung 3.4 sowie die Tabellen 3.1 und 3.2)

- 1) Alle Schätzergebnisse beruhen auf der Annahme, dass der Effekt von PA auf LEZU und der Effekt von SEX2 auf LEZU in allen Ländern gleich ist. Dies ist sehr wahrscheinlich eine falsche Annahme, die hier nur getroffen wird, um systematisch in die MLA einzuführen.
- 2) Die Stärke des Effekts von PA auf LEZU beträgt 0,348 und ist hochsignifikant. Wenn sich PA um eine empirische Einheit erhöht, steigt der zu erwartende LEZU-Wert um 0,35 Punkte an (LEZU-Skala: 1,00-7,00).
- 3) Die Stärke des Effekts von SEX2 auf LEZU beträgt 0,122 und ist hochsignifikant. Wenn sich SEX2 um eine empirische Einheit erhöht d.h. wenn die Analyse von der Gruppe der männlichen Befragten zu der Gruppe der weiblichen Befragten wechselt, steigt der zu erwartende LEZU-Wert um 0,12 Punkte an (LEZU-Skala: 1,00-7,00).
- 4) Die Effekte von PA und SEX2 basieren auf unterschiedlich skalierten Variablen und können deshalb nicht direkt miteinander verglichen werden. Sie könnten miteinander verglichen werden, wenn PA dichotomisiert würde (z.B. mittels Mediansplit) (vgl. dazu das folgende Modell 08). Möglich wäre es, unterschiedliche Signifikanzniveaus miteinander zu vergleichen. Allerdings ist dies hier nicht möglich, weil für jeden Effekt gilt: "hochsignifikant".
- 5) Der durchschnittliche LEZU-Wert beträgt 4,35, wenn PA und SEX2 gleich null sind. Da PA zentriert wurde, ist PA dann gleich null, wenn Befragte den Durchschnittswert von PA aufweisen. Für SEX2 gilt ein Wert von 0, wenn die Gruppe der männlichen Befragten betrachtet wird.
- 6) Die Varianz des L2-Intercepts ist beträchtlich und hochsignifikant (vgl. Tab. 3.2). Das Intercept variiert deutlich von Land zu Land. Ähnliches kann man von den PA- und SEX2-Slopes nicht sagen, denn diese sind ja auf einen festen Wert fixiert. Allerdings signalisieren die AIC/BIC-Werte von Modell 03, die kleiner sind als diejenigen von Modell 02, dass schon eine Erweiterung des Modells um einen weiteren L1-Prädiktor zu einer Verbesserung der Anpassung des Modells führt, auch wenn die länderspezifischen PA-Slopes fixiert sind.
- 7) Viele Modellannahmen multipler OLS-Regressionsschätzungen, die auch für eine ML-Schätzung des Modells 03 gelten sollten (z.B. Abwesenheit von Multikollinearität, vgl. Kap. A2, Pkt. 7) wurden hier nicht untersucht (was allerdings üblicherweise gemacht werden müsste).



Tab. 3.3: Mplus-Inputfile für Modell 03

```
TITLE:
  Modell "MLA 03"
  mit L1-Prädiktor PA (zentriert mit Grand Mean) und mit L1-Prädiktor SEX2 (dichotom 0/1=m/w)
  ohne L2-Prädiktoren
  mit random intercepts und mit fixed slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE nation lezu pa sex2;
  ! nation ist die Cluster-Variable
  ! lezu ist die abh. Y-Variable
  ! pa ist X1-Prädiktor
  ! sex2 ist X2-Prädiktor

WITHIN ARE pa sex2;
  ! hiermit werden PA und SEX2 als L1-Prädiktoren definiert

CLUSTER=nation;

DEFINE:
  IF(sex EQ 1) THEN sex2=0; ! 0=m
  IF(sex EQ 2) THEN sex2=1; ! 1=w

  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert
  ! (keine Zentrierung bei Dummies)

ANALYSIS:
  TYPE = TWOLEVEL;

MODEL: %WITHIN%
  lezu ON pa sex2;

OUTPUT: sampstat;
```

Tab. 3.4: Mplus-Ausgabe von Modellschätzung 03 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters		5		
Loglikelihood				
H0 Value		-9910.468		
H0 Scaling Correction Factor for MLR		2.3056		
H1 Value		-9910.468		
H1 Scaling Correction Factor for MLR		2.3056		
Information Criteria				
Akaike (AIC)		19830.937		
Bayesian (BIC)		19865.154		
Sample-Size Adjusted BIC		19849.265		
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
LEZU	ON			
PA	0.348	0.018	19.328	0.000
SEX2	0.122	0.034	3.635	0.000
Residual Variances				
LEZU	1.003	0.034	29.571	0.000
Between Level				
Means				
LEZU	4.350	0.075	58.247	0.000
Variances				
LEZU	0.179	0.039	4.645	0.000



Modell 04:

Zwei-Ebenen-Modell,

Random-Intercepts-Modell

ohne L1-Prädiktoren

mit L2-Prädiktor "GDP" (zur Schätzung von Intercepts/Mean)  
("intercepts-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + e_{ij}$  (4.1)

L2-Gleichung:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (4.2a)

"fixed part":  $\gamma_{00}, \gamma_{01}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j})$

(Das Modell 04 unterscheidet sich dadurch von Modell 01, dass nunmehr ein Z-Prädiktor (GDP=gross domestic product) in das Modell aufgenommen wird. Das Modell 04 entspricht dem Modell 5.4.3 und dessen Formalisierung in: Geiser 2010: 217- 220).

Abb. 4.1: Graphische Skizze des Gleichungssystems von Modell 04

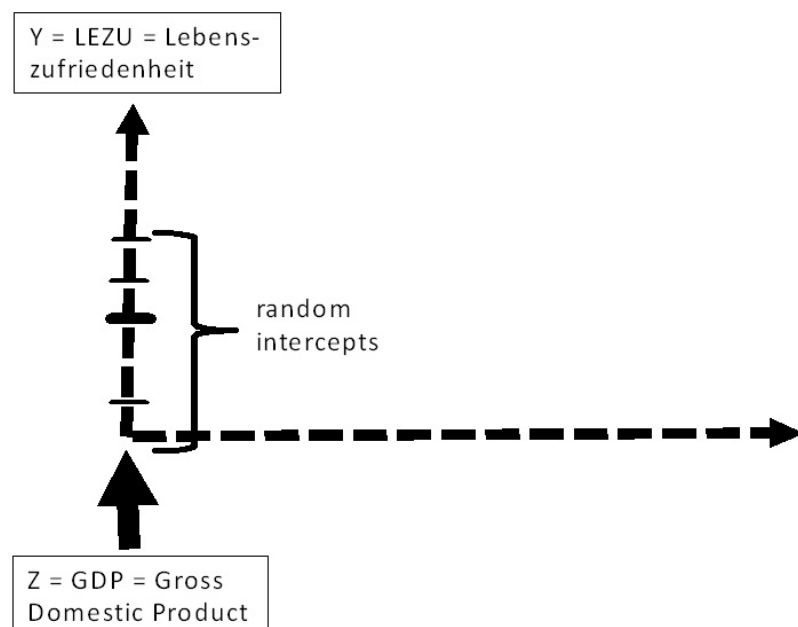
$$Y_{ij} = \beta_{0j} + e_{ij}$$

↓

$$= \gamma_{00} + \gamma_{01}Z_j + u_{0j}$$

$\beta$ : beta      Y: LEZU  
 $\gamma$ : gamma    Z: GDP

Abb. 4.2: Stilisierte Darstellung der Modell-Logik von Modell 04



Zur Logik von Modell 04:

Im Modell 04 gibt es nur einen Prädiktor auf der Makroebene (Z-Prädiktor). Dieser Z-Prädiktor ist in unserem Beispiel das Bruttoinlandsprodukt (GDP = gross domestic product) eines jeden der 40 untersuchten Länder/Staaten/Cluster. Der Z-Prädiktor soll die länderspezifischen Means von  $Y=LEZU$  mittels Linearschätzung ermitteln. Mit dem Z-Prädiktor sollen mithin die länderspezifischen Unterschiede in den Mittelwerten der abhängigen Variablen (hier:  $Y=LEZU$ ) erklärt werden. Mit der Modellschätzung soll sodann ermittelt werden:

- a) wie hoch der mittlere LEZU-Wert in jedem Land ist, wenn die Schätzung nur mit einem L2-Prädiktor (GDP) und ohne L1-Prädiktoren erfolgt,
- b) wie hoch der Effekt von GDP auf LEZU im bivariaten Zwei-Ebenen-Modell ist.

Die Abbildung 4.2 verdeutlicht die Logik von Modell 04. Die Abbildung zeigt ein zweidimensionales Koordinatensystem (gestrichelte Linien). Auf der senkrechten Koordinate werden die geschätzten Werte für  $Y/LEZU$  abgetragen. Auf der waagrechten Koordinate könnten die  $X$ -Werte für L1-Prädiktoren abgetragen werden, die es aber im Modell nicht gibt. Deshalb bleibt die waagrechte  $X$ -Koordinate in Abbildung 4.2 "leer".

Auf der  $Y$ -Koordinate in Abb. 4.2 markieren kleine waagrechte Striche die Lage von Intercepts der Gl. 4.1, die für die verschiedenen untersuchten Länder unterschiedlich groß ausfallen können (random intercepts). Die Ausprägungen dieser länderspezifischen Intercepts sollen in Modell 04 mittels des L2-Prädiktors "Z/GDP" geschätzt werden. Dies soll in der Abbildung durch den sehr fett eingezeichneten Pfeil, der "von unten" auf die  $Y$ -Koordinate gerichtet ist, symbolisiert werden.

Zur Interpretation des Gleichungssystems von Modell 04:

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (4.1)$$

$Y$  abh. Variable LEZU auf der Mikroebene L1

$\beta_{0j}$  Für jeden Befragten wird ein LEZU-Wert geschätzt, der identisch mit dem LEZU-Mittelwert aller Befragten eines jeweiligen Staates/Landes ist. Dieser Mittelwert ist identisch mit dem Achsenabschnitt  $\beta_{0j}$ .

$e_{ij}$  Der mit  $\beta_{0j}$  geschätzte  $Y$ -Wert für jeden Befragten innerhalb eines bestimmten Landes entspricht oftmals nicht dessen tatsächlichen  $Y$ -Wert. Der reale Wert wird oberhalb oder unterhalb des geschätzten Achsenabschnitts liegen. Die Differenz zwischen geschätztem und empirischem  $Y$ -Wert wird als "e" bezeichnet.

$\text{Var}(e_{ij})$  Varianz der Residualvariable  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster/Staaten/Länder). Die Varianz betrifft die Streuung der Abweichungen der individuellen LEZU-Werte vom jeweiligen länderspezifischen LEZU-Mittelwert, wenn der Einfluss von  $Z$  kontrolliert wird (bzw. geschätzt wird).

Mplus: Within Level / Variances LEZU = 1,229 (p=0,000)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (4.2a)$$

$\gamma_{00}$  Für die Lineargleichung zur Schätzung der länderspezifischen LEZU-Mittelwerte wird ein Y-Achsenabschnitt geschätzt ( $\gamma_{00}$ ), der für alle Länder gleich ist und identisch mit dem Grand Mean von LEZU ist. Der Achsenabschnitt ( $\gamma_{00}$ ) ist somit ein fester Parameter (fixed coefficient), der nicht variieren kann.

Mplus: Between Level / Intercepts LEZU = 3,150 (p=0,000)

$\gamma_{01}$  ist der Slope in der Lineargleichung zur Schätzung der länderspezifischen LEZU-Mittelwerte durch den Effekt des länderspezifischen L2-Prädiktors GDP (Vorsicht: vgl. Interpretationshinweis 1)

Mplus: Between Level / LEZU ON GDP = 1,715 (p=0,000)

Z L2-Prädiktor GDP

$u_{0j}$  Die Residualgröße "u<sub>0j</sub>" bezeichnet die Differenz zwischen tatsächlichem und geschätztem LEZU-Mittelwert.

Var( $u_{0j}$ ) Varianz von  $u_{0j}$  (s.o.); Varianz von länderspezifischen LEZU-Mittelwerten, die nicht durch Z vorhergesagt werden kann

Mplus: Between Level / Residual Variances LEZU = 0,198 (p=0,000)

#### Zur Interpretation der Schätzung von Modell 04

(vgl. dazu auch die folgende Abbildung 4.3 sowie die Tabellen 4.1 und 4.2)

- 1) Die Größe des geschätzten Slope-Koeffizienten von GDP beträgt 1,715 und ist hochsignifikant. Der Wert bedeutet, dass für jeden Anstieg von GDP (Skala: 0,36 ... 0,94) um eine Messeinheit (= 0,1) die Lebenszufriedenheit (LEZU-Skala: 1 bis 7) um  $1,715 \cdot 0,1 = 0,1715$  Punkte ansteigt (würde GDP mit dem Zehnfachen der ursprünglichen GDP-Skala gemessen (3,6 ... 9,4), so wäre eine Messeinheit gleich 1,0 und der Slope-Koeffizient wäre 0,1715).
- 2) Die Varianz des geschätzten GDP-Slope-Koeffizienten ist hochsignifikant (vgl. Tab. 4.2). Sie indiziert, dass der GDP-Effekt deutlich zwischen den einzelnen Ländern/Staaten variiert und dass diese Variation noch so viel ungebundene L2-Varianz erzeugt, dass in der Modellschätzung auch noch die Effekte weiterer Z-Variablen untersucht werden können
- 3) Das Intercept auf der L2-Ebene wird als 3.150 geschätzt. Das ist der innerhalb des spezifizierten Modells geschätzte Mittelwert von LEZU, wenn GDP unberücksichtigt bleibt bzw. gleich 0.00 ist.

Abb. 4.3: Graphische Skizze des Gleichungssystems von Modell 04 mit Schätzergebnaten

$$\begin{aligned}
 Y_{ij} &= \beta_{0j} + e_{ij} \\
 &\downarrow \\
 &= \gamma_{00} + \gamma_{01}(\text{GDP}_j) + u_{0j} \\
 &(3,15^{***}) + (1,71^{***})(\text{GDP}_j) + u_{0j} \\
 &(3,15^{***}) + (0,17^{***})(\text{GDP}_j \times 10) + u_{0j}
 \end{aligned}$$

β: beta    γ: LEZU  
 γ: gamma    Z: GDP

Tab. 4.1: Ergebnisse der MLA-Schätzung von Modell 04: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept (γ <sub>00</sub> ) LEZU	3,150	0,000
L2-Slope (γ <sub>01</sub> ) LEZU ON GDP	1,715	0,000

Tab. 4.2: Ergebnisse der MLA-Schätzung von Modell 04: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Varianz, LEZU, Var(e <sub>ij</sub> )	1,229	0,000
L2-Residualvarianz, LEZU, Var(u <sub>0j</sub> )	0,198	0,000

Tab. 4.3: Mplus-Inputfile für Modell 04

```
TITLE:
  Modell "MLA 04"
  mit einem L2-Prädiktor "gdp"
  mit random intercepts

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
            KLA_F  LEZU  PA  NA  BAL  IC
            GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
            PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE nation lezu gdp;
  ! nation ist die Cluster-Variable
  ! lezu ist die abh. Y-Variable
  ! gdp ist Z-Prädiktor

!WITHIN ARE pa;
  ! da auf L1 nur Random Intercepts sind, braucht man hier kein statement,
  ! die Residualvarianz auf L1 wird automatisch geschätzt

BETWEEN = gdp;
  ! definiert gdp als L2-Variable

CLUSTER=nation;

ANALYSIS:
  TYPE = TWOLEVEL;

MODEL: %BETWEEN%
  lezu ON gdp;
```

Tab. 4.4: Mplus-Ausgabe von Modellschätzung 04 (ausgewählte Bereiche)

MODEL FIT INFORMATION					
Number of Free Parameters		4			
Loglikelihood					
H0 Value		-10609.841			
H0 Scaling Correction Factor for MLR		1.6003			
H1 Value		-10609.841			
H1 Scaling Correction Factor for MLR		1.6003			
Information Criteria					
Akaike (AIC)		21227.683			
Bayesian (BIC)		21255.057			
Sample-Size Adjusted BIC		21242.346			
(n* = (n + 2) / 24)					
MODEL RESULTS					
		Two-Tailed			
	Estimate	S.E.	Est./S.E.	P-Value	
Within Level					
Variances					
	LEZU	1.229	0.040	30.471	0.000
Between Level					
LEZU	ON				
	GDP	1.715	0.428	4.003	0.000
Intercepts					
	LEZU	3.150	0.320	9.852	0.000
Residual Variances					
	LEZU	0.198	0.039	5.126	0.000

Modell 05:

Zwei-Ebenen-Modell,

Random-Intercepts-Modell,

ohne L1-Prädiktoren,

mit L2-Prädiktor "GDP2" (dichotom, Mediansplit)

zur Schätzung von Intercepts/Means  
("intercepts-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + e_{ij}$  (5.1)

L2-Gleichung:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (5.2a)

"fixed part":  $\gamma_{00}, \gamma_{01}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j})$

(Das Modell 05 unterscheidet sich allein dadurch von Modell 04, dass die Z-Variable "GDP2" nunmehr dichotomisiert ist (Mediansplit). Das Modell 05 entspricht dem Modell 5.4.3 und dessen Formalisierung in: Geiser 2010: 217 - 220).

Abb. 5.1: Graphische Skizze des Gleichungssystems von Modell 05

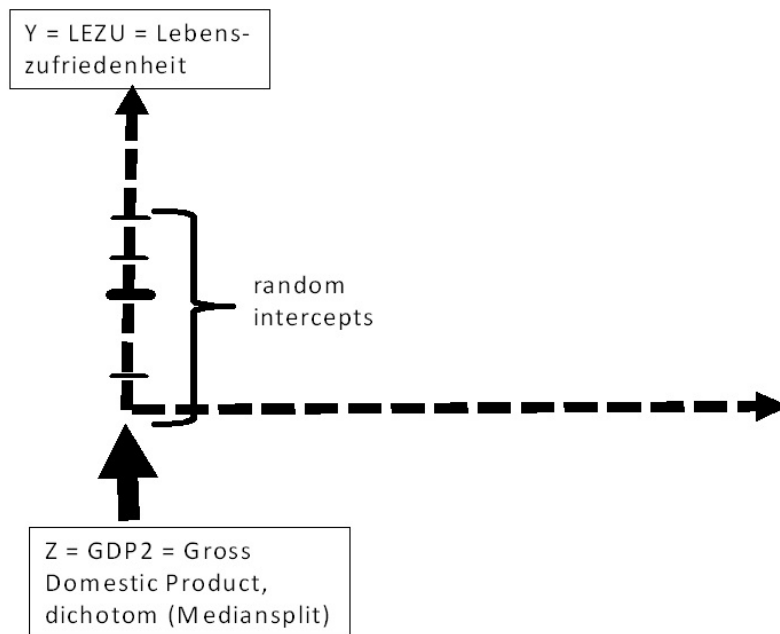
$$Y_{ij} = \beta_{0j} + e_{ij}$$

↓

$$= \gamma_{00} + \gamma_{01}Z_j + u_{0j}$$

$\beta$ : beta      Y: LEZU  
 $\gamma$ : gamma    Z: GDP2

Abb. 5.2: Stilisierte Darstellung der Modell-Logik von Modell 05



Zur Logik von Modell 05:

Im Modell 05 gibt es nur einen Prädiktor auf der Makroebene (Z-Prädiktor). Dieser Z-Prädiktor ist in unserem Beispiel das dichotomisierte Bruttoinlandsprodukt ( $GDP2 = \text{gross domestic product}$ ) eines jeden der 40 untersuchten Länder/Staaten/Cluster (Mediansplit). Der Z-Prädiktor soll die länderspezifischen Means von  $Y=LEZU$  mittels Linearschätzung ermitteln. Mit dem Z-Prädiktor sollen mithin die länderspezifischen Unterschiede in den Mittelwerten der abhängigen Variablen (hier:  $Y=LEZU$ ) erklärt werden. Mit der Modellschätzung soll sodann ermittelt werden:

- a) wie hoch der mittlere LEZU-Wert in jedem Land ist, wenn die Schätzung nur mit einem dichotomen L2-Prädiktor ( $GDP2$ ) und ohne L1-Prädiktoren erfolgt,
- b) wie hoch der Effekt von  $GDP2$  auf  $LEZU$  im bivariaten Zwei-Ebenen-Modell ist.

Die Abbildung 5.2 verdeutlicht die Logik von Modell 05. Die Abbildung zeigt ein zweidimensionales Koordinatensystem (gestrichelte Linien). Auf der senkrechten Koordinate werden die geschätzten Werte für  $Y/LEZU$  abgetragen. Auf der waagrechten Koordinate könnten die X-Werte für L1-Prädiktoren abgetragen werden, die es aber im Modell nicht gibt. Deshalb bleibt die waagrechte X-Koordinate in Abbildung 5.2 "leer".

Auf der Y-Koordinate in Abb. 5.2 markieren kleine waagrechte Striche die Lage von Intercepts der Gl. 5.1, die für die verschiedenen untersuchten Länder unterschiedlich groß ausfallen können (random intercepts). Die Ausprägungen dieser länderspezifischen Intercepts sollen in Modell 05 mittels des L2-Prädiktors "Z/ $GDP2$ " geschätzt werden. Dies soll in der Abbildung durch den sehr fett eingezeichneten Pfeil, der "von unten" auf die Y-Koordinate gerichtet ist, symbolisiert werden.

Zur Interpretation des Gleichungssystems von Modell 05:

Mplus liefert die folgenden Schätzwerte für die freien Parameter von Modell 05 (zur Erläuterung der hier nicht beschriebenen Gleichungsterme vgl. die Ausführungen zu Modell 04):

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (5.1)$$

$\text{Var}(e_{ij})$  Varianz der Residualvariable  $e_{ij}$  auf der L1-Ebene.

Mplus: Within Level / Variances  $LEZU = 1,229$  ( $p=0,000$ )

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (5.2a)$$

$\gamma_{00}$  Achsenabschnitt zur Schätzung von  $\beta_{0j}$  (ist für alle Länder gleich: Grand Mean von  $LEZU$ )

Mplus: Between Level / Intercepts  $LEZU = 4,137$  ( $p=0,000$ )

$\gamma_{01}$  ist der  $GDP2$ -Slope zur Schätzung der länderspezifischen  $LEZU$ -Mittelwerte

Mplus: Between Level /  $LEZU$  ON  $GDP2 = 0,521$  ( $p=0,000$ )

$\text{Var}(u_{0j})$  Varianz von  $u_{0j}$  (s.o.)

Mplus: Between Level / Residual Variances  $LEZU = 0,212$  ( $p=0,000$ )



Zur Interpretation der Schätzung von Modell 05:

- 1) alle GDP2-Effekte sind als Gruppenunterschiede zwischen einerseits den Ländern mit unterdurchschnittlichem GDP und andererseits den Ländern mit überdurchschnittlichem GDP zu interpretieren (GDP wurde mittels Mediansplit dichotomisiert und ist nunmehr: GDP2);
- 2) Die AIC- und BIC-Werte (vgl. Kap. B5) von Modell 05 sind deutlicher schlechter (modelllogisch: höher) als die Werte von Modell 04. Dies ist eine Folge des Informationsverlusts in Modell 05 aufgrund der dichotomen Reskalierung von GDP.
- 3) Die Größe des geschätzten Slope-Koeffizienten von GDP2 beträgt 0,521 und ist hochsignifikant. Der Wert bedeutet, dass für den Anstieg von GDP2 (Skala: 0/1) um eine Messeinheit (= 1,00) die Lebenszufriedenheit (LEZU-Skala: 1 bis 7) um  $0,521 \times 1,0 = 0,521$  Punkte ansteigt.
- 4) Die Varianz des geschätzten GDP2-Slope-Koeffizienten ist hochsignifikant (vgl. Tab. 5.2). Sie indiziert, dass der GDP2-Effekt deutlich zwischen den einzelnen Ländern/Staaten variiert und dass diese Variation noch so viel ungebundene L2-Varianz erzeugt, dass in der Modellschätzung auch noch die Effekte weiterer Z-Variablen untersucht werden können
- 5) Das Intercept auf der L2-Ebene wird als 4,137 geschätzt. Das ist der innerhalb des spezifizierten Modells geschätzte Mittelwert von LEZU, wenn GDP2 unberücksichtigt bleibt bzw. gleich 0.00 ist.

**Abb. 5.3:** Graphische Skizze des Gleichungssystems von Modell 05 mit Schätzergebnaten

$$\text{LEZU}_{ij} = \beta_{0j} + e_{ij}$$

↓

$$= \gamma_{00} + \gamma_{01}(\text{GDP2}_{ij}) + u_{0j}$$

(4,13)\*\*\* (0,52\*\*\*)

β: beta    Y: LEZU  
γ: gamma   Z: GDP2

**Tab. 5.1:** Ergebnisse der MLA-Schätzung von Modell 05: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	4,137	0,000
L2-Slope, ( $\gamma_{10}$ ) LEZU ON GDP2	0,521	0,000

**Tab. 5.2:** Ergebnisse der MLA-Schätzung von Modell 05: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Varianz, LEZU, Var( $e_{ij}$ )	1,229	0.000
L2-Residualvarianz, LEZU Var( $u_{0j}$ )	0,212	0,000

Tab. 5.3: Mplus-Inputfile für Modell 05

```

TITLE:
  Modell "MLA 05"
  mit einem L2-Prädiktor "gdp", der nunmehr dichotomisiert ist
  mit random intercepts
  mit fixed slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE nation lezu gdp2;
  ! nation ist die Cluster-Variable
  ! lezu ist die abh. Y-Variable
  ! gdp2 ist der dichotomisierte Z-Prädiktor

!WITHIN ARE pa;
  ! da auf L1 nur Random Intercepts sind, braucht man hier kein statement
  ! Residualvarianz auf L1 wird automatisch geschätzt

BETWEEN = gdp2;
  ! definiert gdp als L2-Variable

CLUSTER=nation;

! hier kommt Mediansplit:
DEFINE:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74) THEN gdp2=1;

ANALYSIS:
  TYPE = TWOLEVEL;

MODEL: %BETWEEN%
  lezu ON gdp2;

OUTPUT:
  sampstat;
  !stdyx; !verlangt vollstandardisierte Lösung = ist bei dichotomen Prädiktoren schlecht !!!

```

Tab. 5.4: Mplus-Ausgabe von Modellschätzung 05 (ausgewählte Bereiche)

MODEL FIT INFORMATION					
Number of Free Parameters				4	
Loglikelihood					
H0 Value				-10611.201	
H0 Scaling Correction Factor for MLR			1.6143		
H1 Value				-10611.201	
H1 Scaling Correction Factor for MLR			1.6143		
Information Criteria					
Akaike (AIC)				21230.401	
Bayesian (BIC)				21257.775	
Sample-Size Adjusted BIC				21245.064	
(n* = (n + 2) / 24)					
MODEL RESULTS					
			Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value	
Within Level					
Variances					
	LEZU	1.229	0.040	30.472	0.000
Between Level					
LEZU	ON				
	GDP2	0.521	0.149	3.499	0.000
Intercepts					
	LEZU	4.137	0.113	36.608	0.000
Residual Variances					
	LEZU	0.212	0.043	4.974	0.000

Modell 06:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell

mit L1-Prädiktor "PA" (zentriert)

ohne L2-Prädiktoren

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij}$  (6.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + u_{0j}$  (6.2a)

$\beta_{1j} = \gamma_{10} + u_{1j}$  (6.2b)

“fixed part”:  $\gamma_{00}, \gamma_{10}$

“random part”:  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$

“covariance”:  $\text{Cov}(u_{0j}, u_{1j})$

(Das Modell 06 entspricht dem Modell 5.5.1 und dessen Formalisierung in: Geiser 2010: 223 - 227).

Abb. 6.1: Graphische Skizze des Gleichungssystems von Modell 06

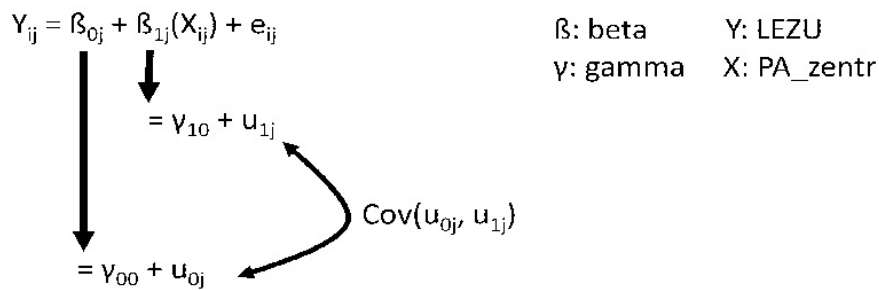
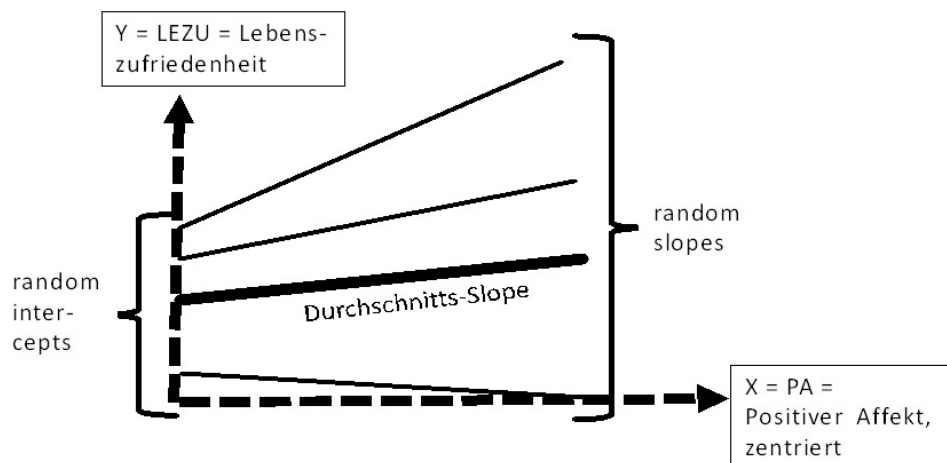


Abb. 6.2: Stilisierte Darstellung der Modell-Logik von Modell 06



Zur Logik von Modell 06:

Die zuvor erläuterten Mehrebenen-Modelle betreffen zumeist eine L1-Regression mit einem variierenden Intercept (bzw. Y-Achsenabschnitt) auf der L2-Ebene. So gab es z.B. einen PA-Effekt auf die indiv. Lebenszufriedenheit, dessen Niveau zwischen versch. Ländern variiert.

Im Folgenden wird ein Modell vorgestellt (Modell 06), bei dem zusätzlich zum variierenden Intercept auch der Slope bzw. die Steigung der Regressionsgeraden zur Prognose von LEZU über die Länder/ Staaten/ Cluster hinweg variiert.

Allerdings enthält Modell 06 keine L2-Prädiktoren zur Erklärung der Varianzen von Intercept und Slope. Im Modell soll lediglich ein L1-Prädiktoreffekt geschätzt werden, dies aber unter Berücksichtigung der Mehrebenenstruktur der Daten. Denn ohne Berücksichtigung der Zwei-Ebenen-Datenstruktur könnten die Standardfehler der Regressionsschätzung unterschätzt werden und es könnten sich verzerrte Signifikanztestes und verzerrte Konfidenzintervalle ergeben, sodass evtl. falsche signifikante Ergebnisse ermittelt würden (nach Geiser 2010: 223f).

Bezogen auf unser LEZU-Beispiel bedeutet Modell 06, dass der PA-Effekt in jedem Land ein anderer sein kann, weil in jedem Land die individuelle Lebenszufriedenheit in unterschiedlicher Weise mit individuell gestaltbaren Lebensbedingungen verknüpft wird. Dann gäbe es einen länderspezifischen PA-Effekt und in jedem Land würde eine andere Regressionsgerade mit anderem Intercept und anderem Slope zur Prognose von LEZU geschätzt werden. Die MLA hätte dann ein mittleres Intercept und einen mittleren Slope für den PA-Effekt zu ermitteln.

Die Abbildung 6.2 verdeutlicht die Logik von Modell 06. Das Koordinatensystem (gestrichelte Linien) besteht aus einer senkrechten Y-Koordinate und eine waagrechten X-Koordinate für die beiden Variablen "LEZU" und "PA". Die Variable LEZU ist die abh. Variable und die Variable PA ist die unabh. Variable (bzw. der Prädiktor). Im Koordinatensystem sind die Regressionsgeraden eingezeichnet (Y wird auf X regrediert). Für jedes untersuchte Land/Cluster wird eine Regressionsgerade geschätzt. Die Regressionsgeraden haben je nach Land unterschiedliche Intercepts (random intercepts) und unterschiedliche Steigungen (random slopes). Die MLA berechnet über alle Länder/Cluster das durchschnittliche Intercept und die durchschnittliche Steigung (den Durchschnitts-Slope).

Zur Interpretation des Gleichungssystems von Modell 06:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (6.1)$$

Y abh. L1-Variable "LEZU=Lebenszufriedenheit"

$\beta_{0j}$  L1-Intercept in der Lineargleichung zur Schätzung von Y mittels Prädiktor "PA=Positiver Affekt (zentriert)" (s.u.). Ist bei PA=0 (=Mittelwert von PA aufgrund der Zentrierung) identisch mit dem durchschnittlichen PA-Wert über alle Länder (s.u. " $\gamma_{00}$ ").

$\beta_{1j}$  PA-Slope in der Lineargleichung zur Schätzung von Y mittels Prädiktor "PA=Positiver Affekt (zentriert)" (s.u. " $\gamma_{10}$ ")

X unabh. L1-Prädiktor "PA=Positiver Affekt (zentriert)"

$e_{ij}$  Differenz zwischen geschätzten und empirischen Y-Werten (L1-Residualvariable).

Var( $e_{ij}$ ) Varianz der Residualvar.  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster). Die Varianz betrifft die Streuung der Abweichungen der indiv. LEZU-Werte vom jeweiligen länderspez LEZU-Mittelwert.

Mplus: Within Level / Variances LEZU = 0,995 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (6.2a)$$

$\gamma_{00}$  Durchschnitt (Grand Mean) von allen länderspezifischen Intercepts. Mittleres Intercept über alle Länder hinweg.

Mplus: Between Level / Means LEZU = 4,415 (p=0,000)

$u_{0j}$  Abweichungen der länderspezifischen Intercepts vom Durchschnitt aller Intercepts (L2-Residualvariable)

Var( $u_{0j}$ ) Varianz von  $u_{0j}$

Mplus: Between Level / Variances LEZU = 0,187 (p=0,000)

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (6.2b)$$

$\gamma_{10}$  Durchschnitt von allen länderspezifischen PA-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA1J = 0,348 (p=0,000)

$u_{1j}$  Differenzen der länderspezifischen PA-Slopes vom Durchschnittswert aller PA-Slopes (L2-Residualvariable)

Var( $u_{1j}$ ) Varianz von  $u_{1j}$

Mplus: Between Level / Variances BETA1J = 0,006 (p=0,000)

Cov( $u_{0j}, u_{1j}$ ) Kovarianz zwischen den länderspezifischen Intercepts und Slopes

Mplus: Between Level / LEZU WITH BETA1J = 0,013 (p=0,035)

#### Zur Interpretation der Schätzung von Modell 06:

(vgl. dazu auch die folgende Abbildung 6.3 sowie die Tabellen 6.1 bis 6.3)

- 1) Im Beispiel wird "PA(zentr) → LEZU" unter Berücksichtigung der Clusterstruktur der Daten mit Mplus geschätzt. Dazu werden von Mplus die "Means" von LEZU und BETA1J ausgegeben. Sie betragen 4,415 und 0,348 (beide hochsignifikant). Sie betreffen a) das mittlere Intercept von LEZU (wenn PA=0) und b) den mittleren Slope über alle Cluster. Aus diesen Werten ergibt sich folgende Modellschätzung (unter Berücksichtigung der Clusterstruktur der Daten mit korrekten Standardfehlern): LEZU = 4,415 + 0,348 (PA zentriert) + e.
- 2) Die L1-Residual-Varianz beträgt 0,995 und ist hochsignifikant (p=0,000). Es können/sollten also noch weitere L1-Prädiktoren in Modell 06 aufgenommen werden.
- 3) Die Varianzen der Koeffizienten beider Schätzgleichungen (6.2a und 6.2b) sind hochsignifikant (p=0,000). Die geschätzten Intercepts und die geschätzten PA-Slopes variieren also nach Staaten- bzw. nach Länderzugehörigkeit.
- 4) Die Kovarianz zwischen den länderspezifischen Intercepts und PA-Slopes beträgt 0,013 und ist signifikant (p=0,035). Bei einem hohen Intercept ist also auch der Slope hoch und umgekehrt.

**Abb. 6.3:** Graphische Skizze des Gleichungssystems von Modell 06 mit Schätzergebnaten

$$\text{LEZU}_{ij} = \beta_{0j} + \beta_{1j}(\text{PA\_zentr}_{ij}) + e_{ij}$$

β: beta  
γ: gamma

$= \gamma_{10} + u_{1j}$   
 (0,34\*\*\*)  
 $= \gamma_{00} + u_{0j}$   
 (4,41\*\*\*)  
 0,013\*

**Tab. 6.1:** Ergebnisse der MLA-Schätzung von Modell 06: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) Mean (LEZU)	4,415	0,000
L2-Slope, ( $\gamma_{10}$ ) Mean, Slope(PA_zentr)	0,348	0,000

**Tab. 6.2:** Ergebnisse der MLA-Schätzung von Modell 06: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	0,995	0,000
L2-Varianz, LEZU, Var( $u_{0j}$ )	0,187	0,000
L2-Varianz, Slope(PA_zentr), Var( $u_{1j}$ )	0,006	0,008

**Tab. 6.3:** Ergebnisse der MLA-Schätzung von Modell 6: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}$ , $u_{1j}$ )	0,013	0,035



Tab. 6.4: Mplus-Inputfile für Modell 06

```

TITLE:
  Modell "MLA 06"
  mit einem L1-Prädiktor "pa"
  mit keinem L2-Prädiktor
  mit random intercept
  mit random slope

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa;  ! pa ist X-Var: positiver Affekr

WITHIN=pa; ! pa wird als L1-Prädiktor definiert

CLUSTER=nation;

DEFINE:
  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
    ! wenn random slopes geschätzt werden sollen

MODEL: %WITHIN%
  Beta1j | lezu ON pa; ! Spezifikation der L1-Regression mit Random Slope
    ! Beta1j ist Name für Random Slope

%BETWEEN%
  lezu WITH Beta1j; ! Kovarianz zwischen Random Intercept und random slope

OUTPUT:
  sampstat;

```

Tab. 6.5: Mplus-Ausgabe von Modellschätzung 06 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters	6			
Loglikelihood				
H0 Value	-9899.465			
H0 Scaling Correction Factor for MLR	1.3715			
Information Criteria				
Akaike (AIC)	19810.931			
Bayesian (BIC)	19851.991			
Sample-Size Adjusted BIC	19832.925			
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	0.995	0.034	29.240	0.000
Between Level				
LEZU WITH				
BETA1J	0.013	0.006	2.111	0.035
Means				
LEZU	4.415	0.070	63.247	0.000
BETA1J	0.348	0.016	21.856	0.000
Variances				
LEZU	0.187	0.041	4.530	0.000
BETA1J	0.006	0.001	4.259	0.000

Modell 07:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell,

mit L1-Prädiktor "PA" (zentriert),

mit L1-Prädiktor "SEX2" (dichotomisiert, 0/1=m/w),

ohne L2-Prädiktoren

$$\text{L1-Gleichung: } Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (7.1)$$

$$\text{L2-Gleichungen: } \beta_{0j} = \gamma_{00} + u_{0j} \quad (7.2a)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (7.2b)$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (7.2c)$$

"fixed part":  $\gamma_{00}, \gamma_{10}, \gamma_{20}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j}), \text{Var}(u_{2j})$

"covariance":  $\text{Cov}(u_{0j}, u_{1j}), \text{Cov}(u_{0j}, u_{2j}), \text{Cov}(u_{1j}, u_{2j})$

(Modell 07 ist analog zu Modell 06 zu betrachten, allerdings jetzt mit dem zusätzlichen dichotomen L1-Prädiktor "SEX2" (0/1 = m/w))

Abb. 7.1: Graphische Skizze des Gleichungssystems von Modell 07

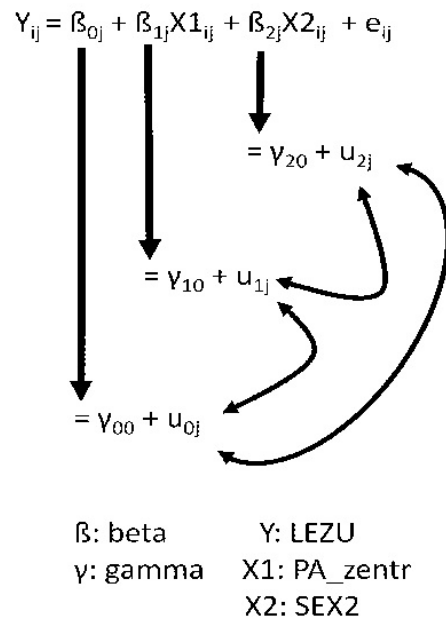
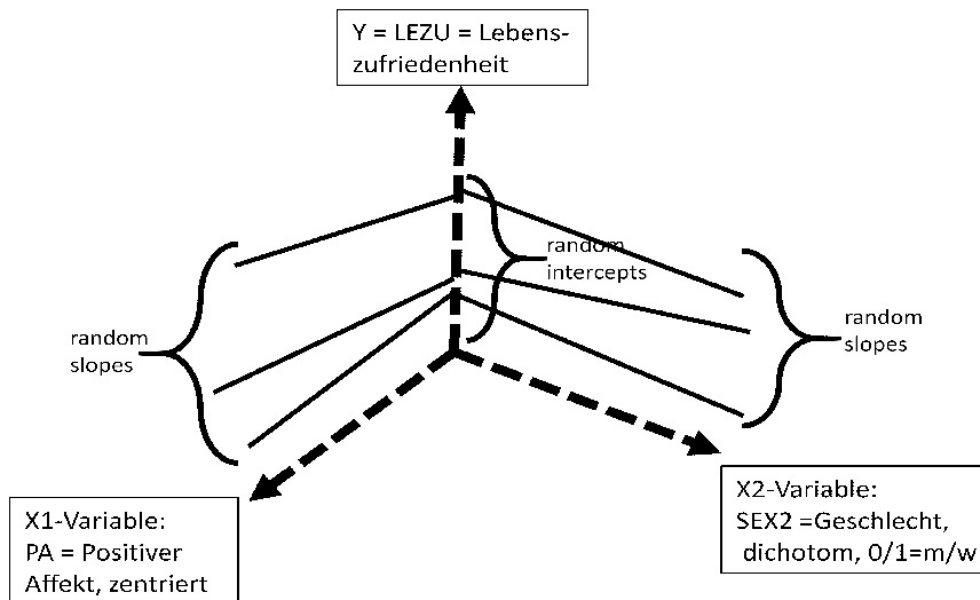


Abb. 7.2: Stilisierte Darstellung der Modell-Logik von Modell 07

Zur Logik von Modell 07:

Die Struktur von Modell 07 ist identisch mit der Struktur von Modell 06. Der große Unterschied zwischen beiden Modellen ist allein die Anzahl der L1-Prädiktoren. In Modell 07 wird als zusätzlicher L1-Prädiktor die Variable “SEX2“ mit den Ausprägungen “0/1=m/w” aufgenommen. Zur weiteren Erläuterung des Modells 07 vgl. die Ausführungen zur Logik von Modell 06.

Die Abbildung 7.2 verdeutlicht die Logik von Modell 07. Das Modell 07 wird im Vergleich zu Modell 06 um eine dritte Koordinate (X2/SEX2) erweitert (vgl. auch “Die Abbildung 6.2 ...”). Die Effekte des neuen Prädiktors SEX2 variieren ebenfalls zwischen den einzelnen Ländern/Clustern, deshalb handelt es sich bei den länderspezifischen Steigungsraten ebenfalls um “random slopes” (wie auch schon beim X1-Prädiktor “PA\_zentr”).

Zur Interpretation des Gleichungssystems von Modell 07:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (7.1)$$

Y abh. L1-Variable “LEZU=Lebenszufriedenheit”

$\beta_{0j}$  L1-Intercept in der Lineargleichung zur Schätzung von LEZU mittels der Prädiktoren “PA=Positiver Affekt (zentriert)” und “SEX2 (dichotom)”. Ist PA=0 (=Mittelwert von PA aufgrund der Zentrierung) und ist gleichzeitig SEX2=0 (=m), so markiert der Koeffizient den durchschnittlichen länderspezifischen LEZU-Wert.

$\beta_{1j}$  partieller PA-Slope in der Lineargleichung zur Schätzung von Y mittels Prädiktor “PA=Positiver Affekt (zentriert)” (s.u. “ $\gamma_{10}$ “)

X1 unabh. L1-Prädiktor “PA=Positiver Affekt (zentriert)”

$\beta_{2j}$  partieller SEX2-Slope in der Lineargleichung zur Schätzung von Y mittels Prädiktor “SEX2 (dichotom)” (s.u. “ $\gamma_{20}$ “)

X2 unabh. L1-Prädiktor “SEX2 (dichotom)”

$e_{ij}$  Differenz zwischen geschätztem und empirischen Y-Wert (L1-Residualvariable)

$\text{Var}(e_{ij})$  Varianz der Residualvariable  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster). Die Varianz betrifft die Streuung der Abweichungen der indiv. LEZU-Werte vom jeweiligen länderspezifischen LEZU-Mittelwert.

Mplus: Within Level / Variances LEZU = 0,989 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (7.2a)$$

$\gamma_{00}$  Durchschnitt (Grand Mean) von allen länderspezifischen Intercepts. Mittleres Intercept über alle Länder hinweg.

Mplus: Between Level / Means LEZU = 4,352 (p=0,000)

$u_{0j}$  Abweichungen der länderspezifischen Intercepts vom Durchschnitt aller Intercepts (L2-Residualvariable)

$\text{Var}(u_{0j})$  Varianz von  $u_{0j}$

Mplus: Between Level / Variances LEZU = 0,205 (p=0,000)

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (7.2b)$$

$\gamma_{10}$  Durchschnitt von allen länderspezifischen partiellen PA-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA1J = 0,346 (p=0,000)

$u_{1j}$  Differenzen der länderspezifischen partiellen PA-Slopes vom Durchschnittswert aller partiellen PA-Slopes. Die Residualvariable macht deutlich, dass der geschätzte PA-Effekt nicht konstant ist, sondern von Land zu Land variiert.

$\text{Var}(u_{1j})$  Varianz von  $u_{1j}$

Mplus: Between Level / Variances BETA1J = 0,006 (p=0,000)

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (7.2c)$$

$\gamma_{20}$  Durchschnitt von allen länderspezifischen partiellen SEX2-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA2J = 0,112 (p=0,000)

$u_{2j}$  Differenzen der länderspezifischen partiellen SEX2-Slopes vom Durchschnittswert aller partiellen SEX2-Slopes. Die Residualvariable macht deutlich, dass der geschätzte SEX2-Effekt nicht konstant ist, sondern von Land zu Land variiert.

$\text{Var}(u_{2j})$  Varianz von  $u_{2j}$

Mplus: Between Level / Variances BETA2J = 0,012 (p=0,167)

$\text{Cov}(u_{0j}, u_{1j})$  Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen PA-Slopes

Mplus: Between Level / LEZU WITH BETA1J = 0,015 (p=0,025)

$\text{Cov}(u_{0j}, u_{2j})$  Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen SEX2-Slopes

Mplus: Between Level / LEZU WITH BETA2J = -0,025 (p=0,171)

$\text{Cov}(u_{1j}, u_{2j})$  Kovarianz zwischen länderspezifischen partiellen PA-Slopes und länderspezifischen SEX2-Slopes

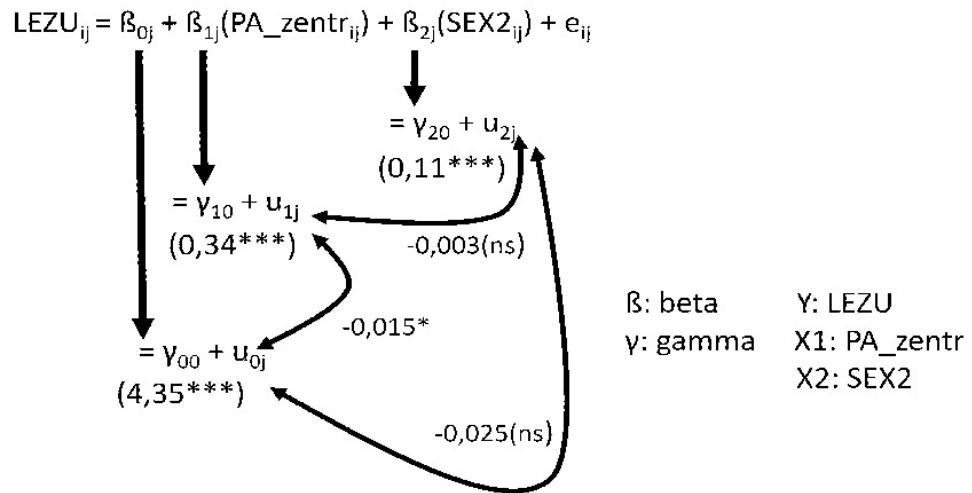
Mplus: Between Level / BETA1J WITH BETA2J = -0,03 (p=0,213)

Zur Interpretation der Schätzung von Modell 07:

(vgl. dazu auch die folgende Abbildung 7.3 sowie die Tabellen 7.1 bis 7.3)

- 1) Im Beispiel wird der partielle mittlere PA-Slope und der partielle mittlere SEX2-Slope unter Berücksichtigung der Clusterstruktur der Daten geschätzt. Die Slopes betragen 0,346 und 0,112. Beide sind hochsignifikant. Die Stärken der Slopes lassen sich nicht direkt miteinander vergleichen, weil sie sehr unterschiedlich skaliert sind. Sie könnten miteinander verglichen werden, wenn PA dichotomisiert wäre (z.B. mittels Mediansplit). Zudem wäre es generell möglich, Slope-Schätzungen hinsichtlich ihrer unterschiedlichen Signifikanzen miteinander zu vergleichen. Allerdings ist dies hier auch nicht möglich, weil für jede Effektschätzung gilt: "hochsignifikant".
- 2) Das mittlere Intercept von LEZU beträgt 4,352 und ist hochsignifikant, wenn PA und SEX2 gleich null sind. Da PA zentriert wurde, ist  $PA=0$ , wenn Befragte den Durchschnittswert von PA aufweisen. Für SEX2 gilt ein Wert von 0, wenn die Gruppe der männlichen Befragten betrachtet wird.
- 3) Die L1-Residual-Varianz beträgt 0,989 und ist hochsignifikant ( $p=0,000$ ). In Modell 01 ohne Residual-Varianz betrug sie noch 1,22. Die L1-Residual-Varianz konnte also mit dem Modell 07 im Vergleich zu Modell 01 um 19% reduziert werden. Auch die AIC/BIC-Werte (vgl. Kap. B5) sind von Modell 01 zu Modell 07 sehr deutlich gefallen (z.B. der AIC-Wert von 21239,129 in M01 auf 19790,669 in M07).
- 4) Die L1-Varianz der Residuen ist hochsignifikant ( $p=0,000$ ). Das Schätzmodell könnte also noch um zusätzliche Prädiktoren erweitert werden, um freie Varianzanteile zu binden.
- 5) Die Varianz der LEZU-Intercepts und die Varianz der PA-Slopes sind hochsignifikant. Beide Größen (LEZU-Intercepts und PA-Slopes) variieren demnach sehr deutlich zwischen den verschiedenen Ländern/Staaten. Für den SEX2-Slope gilt dies nicht. Er variiert zwischen den Ländern/Staaten in nicht-signifikanter Weise.
- 6) Die Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen PA-Slopes beträgt 0,015 und ist signifikant ( $p=0,025$ ). Bei einem hohen Intercept ist also auch der Slope hoch und umgekehrt. Die beiden anderen Kovarianzen sind nicht signifikant.
- 7) Viele Modellannahmen, die für eine ML-Schätzung des Modells 07 gelten sollten (z.B. Abwesenheit von Multikollinearität u.a.) wurden hier aus Zeitgründen nicht untersucht, sollten allerdings üblicherweise bei einer ML-Modellschätzung getestet werden (vgl. dazu Kap. A2).

Abb. 7.3: Graphische Skizze des Gleichungssystems von Modell 07 mit Schätzergebnaten



Tab. 7.1: Ergebnisse der MLA-Schätzung von Modell 07: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) Mean (LEZU)	4,352	0,000
L2-Slope1, ( $\gamma_{10}$ ) Mean, Slope1(PA-zentr)	0,346	0,000
L2-Slope2, ( $\gamma_{20}$ ) Mean, Slope2(SEX2)	0,112	0,000

Tab. 7.2: Ergebnisse der MLA-Schätzung von Modell 07: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	0,989	0,000
L2-Varianz, LEZU, Var( $u_{0j}$ )	0,205	0,000
L2-Varianz, Slope1(PA_zentr) Var( $u_{1j}$ )	0,006	0,008
L2-Varianz, Slope2(SEX2) Var( $u_{2j}$ )	0,012	0,167

Tab. 7.3: Ergebnisse der MLA-Schätzung von Modell 07: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}, u_{1j}$ )	0,015	0,025
Cov( $u_{0j}, u_{2j}$ )	-0,025	0,171
Cov( $u_{1j}, u_{2j}$ )	-0,003	0,21

Tab. 7.4: Mplus-Inputfile für Modell 07

```

TITLE:
  Modell "MLA 07"
  mit zwei L1-Prädiktoren: "pa" und "sex2"
  mit keinem L2-Prädiktor
  mit random intercepts und mit zwei random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
            KLA_F  LEZU  PA  NA  BAL  IC
            GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
            PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa    ! pa ist X-Var: positiver Affekt
  sex2; ! 0/1 = m/w

WITHIN ARE pa sex2 ; ! pa und sex2 werden als L1-Prädiktoren definiert

CLUSTER=nation;

DEFINE:
  IF(sex EQ 1) THEN sex2=0;
  IF(sex EQ 2) THEN sex2=1;

  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
    ! wenn random slopes geschätzt werden sollen

MODEL: %WITHIN%
  Beta1j | lezu ON pa; ! Spezifikation der L1-Regression mit Random Slope
    ! Beta1j ist Name für Random Slope
  Beta2j | lezu ON sex2 ; ! Spezifikation der L1-Regression mit Random Slope
    ! Beta2j ist Name für Random Slope

%BETWEEN%
  lezu WITH Beta1j; ! Kovarianz zwischen Random Intercept und random slope
  lezu WITH Beta2j; ! Kovarianz zwischen Random Intercept und random slope
  Beta1j WITH Beta2j; ! Kovarianz zwischen X1-random slope und X2-random slope

OUTPUT:
  sampstat;

```



Tab. 7.5: Mplus-Ausgabe von Modellschätzung 07 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				10
Loglikelihood				
H0 Value				-9885.335
H0 Scaling Correction Factor for MLR				1.2245
Information Criteria				
Akaike (AIC)				19790.669
Bayesian (BIC)				19859.104
Sample-Size Adjusted BIC				19827.326
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	0.989	0.034	28.695	0.000
Between Level				
LEZU WITH				
BETA1J	0.015	0.007	2.239	0.025
BETA2J	-0.025	0.019	-1.369	0.171
BETA1J WITH				
BETA2J	-0.003	0.003	-1.245	0.213
Means				
LEZU	4.352	0.074	58.958	0.000
BETA1J	0.346	0.016	21.579	0.000
BETA2J	0.112	0.031	3.592	0.000
Variances				
LEZU	0.205	0.049	4.188	0.000
BETA1J	0.006	0.002	4.242	0.000
BETA2J	0.012	0.009	1.381	0.167

Modell 08:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell

mit L1-Prädiktor "PA2" (dichotomisiert, Mediansplit)

mit L1-Prädiktor "SEX2" (dichotomisiert, 0/1=m/w)

ohne L2-Prädiktoren

L1-Gleichung  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (8.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + u_{0j}$  (8.2a)

$\beta_{1j} = \gamma_{10} + u_{1j}$  (8.2b)

$\beta_{2j} = \gamma_{20} + u_{2j}$  (8.2c)

“fixed part”:  $\gamma_{00}, \gamma_{10}, \gamma_{20}$

“random part”:  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j}), \text{Var}(u_{2j})$

“covariances”:  $\text{Cov}(u_{0j}, u_{1j}), \text{Cov}(u_{0j}, u_{2j}), \text{Cov}(u_{1j}, u_{2j})$

(Das Modell 08 ist analog zu Modell 07 konstruiert, allerdings jetzt mit einem dichotomen L1-Prädiktor "PA2" (Mediansplit).)

Abb. 8.1: Graphische Skizze des Gleichungssystems von Modell 08

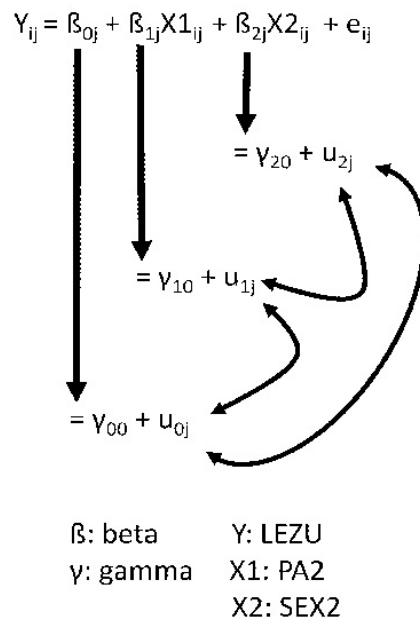
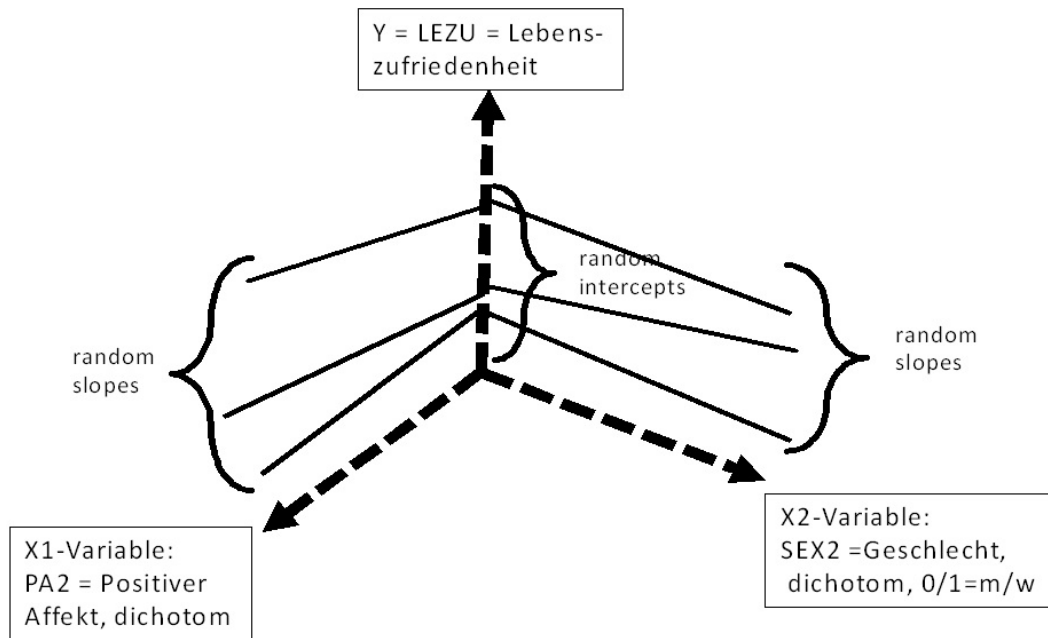


Abb. 8.2: Stilisierte Darstellung der Modell-Logik von Modell 08

Zur Logik von Modell 08:

Das Modell 08 unterscheidet sich allein dadurch von Modell 07, dass der L1-Prädiktor "PA" nunmehr mittels Mediansplit dichotomisiert wurde (neu: "PA2"). Zur Logik von Modell 08 vgl. deshalb die Informationen "Zur Logik von Modell 06" und "Zur Logik von Modell 07".

Die Abbildung 8.2 verdeutlicht die Logik von Modell 08. Nähere Ausführungen dazu finden sich in den Textabschnitten "Die Abbildung 6.2 verdeutlicht..." und "Die Abbildung 7.2 verdeutlicht...".

Zur Interpretation des Gleichungssystems von Modell 08:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (8.1)$$

Y abh. L1-Variable "LEZU=Lebenszufriedenheit"

$\beta_{0j}$  L1-Intercept in der Lineargleichung zur Schätzung von LEZU mittels der Prädiktoren "PA2=Positiver Affekt (dichotom)" und "SEX2 (dichotom)". Für PA2=0 (=unterdurchschnittl. PA-Wert wg. Mediansplit) und (gleichzeitig) SEX2=0 (=m) markiert der Koeffizient einen länderspezifischen LEZU-Wert, der als Referenzwert zur Beurteilung der geschätzten Slopes (s.u.) benutzt werden kann.

$\beta_{1j}$  partieller PA2-Slope in der Lineargleichung zur Schätzung von LEZU mittels Prädiktor "PA2=Positiver Affekt (dichotomisiert)" (s.u. " $\gamma_{10}$ ")

X1 unabh. L1-Prädiktor "PA2=Positiver Affekt (dichotom)"

$\beta_{2j}$  partieller SEX2-Slope in der Lineargleichung zur Schätzung von Y mittels Prädiktor "SEX2 (dichotom)" (s.u. " $\gamma_{20}$ ")

X2 unabh. L1-Prädiktor "SEX2 (dichotom)"

$e_{ij}$  Differenz zwischen geschätztem und empirischem Y-Wert (L1-Residualvariable)

$\text{Var}(e_{ij})$  Varianz der Residualvariable  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster). Die Varianz betrifft die Streuung der Abweichungen der individuellen LEZU-Werte vom jeweiligen länderspez. LEZU-Mittelwert.

Mplus: Within Level / Variances LEZU = 1,084 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (8.2a)$$

$\gamma_{00}$  Durchschnitt (Grand Mean) von allen länderspezifischen Intercepts. Mittleres Intercept über alle Länder hinweg.

Mplus: Between Level / Means LEZU = 3,995 (p=0,000)

$u_{0j}$  Abweichungen der länderspezifischen Intercepts vom Durchschnitt aller Intercepts (L2-Residualvariable)

$\text{Var}(u_{0j})$  Varianz von  $u_{0j}$

Mplus: Between Level / Variances LEZU = 0,216 (p=0,000)

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (8.2b)$$

$\gamma_{10}$  Durchschnitt von allen länderspezifischen partiellen PA2-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA1J = 0,761 (p=0,000)

$u_{1j}$  Differenzen der länderspezifischen partiellen PA2-Slopes vom Durchschnittswert aller partiellen PA2-Slopes. Die Residualvariable macht deutlich, dass der geschätzte PA2-Effekt nicht konstant ist, sondern von Land zu Land variiert.

$\text{Var}(u_{1j})$  Varianz von  $u_{1j}$

Mplus: Between Level / Variances BETA1J = 0,034 (p=0,005)

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (8.2c)$$

$\gamma_{20}$  Durchschnitt von allen länderspezifischen partiellen SEX2-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA2J = 0,138 (p=0,000)

$u_{2j}$  Differenzen der länderspezifischen partiellen SEX2-Slopes vom Durchschnittswert aller partiellen SEX2-Slopes. Die Residualvariable macht deutlich, dass der geschätzte SEX2-Effekt nicht konstant ist, sondern von Land zu Land variiert.

$\text{Var}(u_{2j})$  Varianz von  $u_{2j}$

Mplus: Between Level / Variances BETA2J = 0,012 (p=0,183)

$\text{Cov}(u_{0j}, u_{1j})$  Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen PA2-Slopes

Mplus: Between Level / LEZU WITH BETA1J = 0,003 (p=0,826)

$\text{Cov}(u_{0j}, u_{2j})$  Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen SEX2-Slopes

Mplus: Between Level / LEZU WITH BETA2J = -0,028 (p=0,100)

$\text{Cov}(u_{1j}, u_{2j})$  Kovarianz zwischen länderspezifischen partiellen PA2-Slopes und länderspezifischen SEX2-Slopes

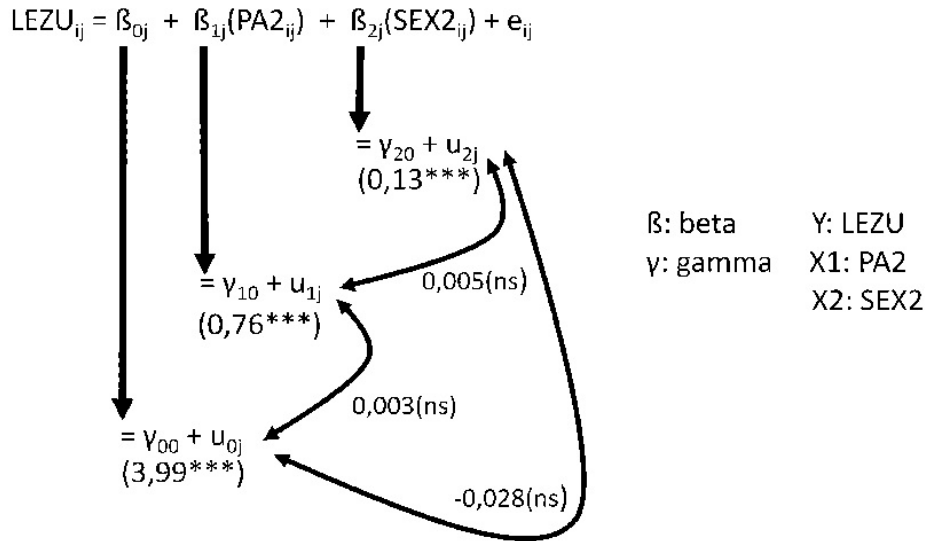
Mplus: Between Level / BETA1J WITH BETA2J = 0,005 (p=0,611)

Zur Interpretation der Schätzung von Modell 08:

(vgl. dazu auch die folgende Abbildung 8.3 sowie die Tabellen 8.1 bis 8.3)

- 1) Im Beispiel wird der partielle mittlere PA2-Slope und der partielle mittlere SEX2-Slope unter Berücksichtigung der Clusterstruktur der Daten geschätzt. Die Slopes betragen 0,761 und 0,138. Beide sind hochsignifikant. Da die beiden Prädiktoren dichotomisiert sind, können die Stärken ihrer Slopes direkt miteinander verglichen werden. Demnach ist der Effekt von PA2 mehr als fünfmal so stark wie der Effekt von SEX2. Ein Wechsel von unterdurchschnittlichem positiven Affekt (PA2=0) zu überdurchschnittlichem positiven Affekt (PA2=1) erhöht den durchschnittlichen LEZU-Wert von 3,995 um 0,761 Punkte. Demgegenüber haben weibliche Befragte (SEX2=1) nur einen im Vergleich zu männlichen Befragten (SEX2=0) erhöhten LEZU-Wert um 0,138 Punkte.
- 2) Das mittlere Intercept von LEZU beträgt 3,995 und ist hochsignifikant, wenn PA2 und SEX2 gleich null sind. Da PA dichotomisiert wurde (Mediansplit), ist PA2=0, wenn Befragte einen unterdurchschnittlichen PA-Wert aufweisen. Für SEX2 gilt ein Wert von 0, wenn die Gruppe der männlichen Befragten betrachtet wird.
- 3) Die L1-Residual-Varianz beträgt 1,084 und ist hochsignifikant ( $p=0,000$ ). In Modell 1 ohne Residual-Varianz betrug sie noch 1,22. Die freie L1-Residual-Varianz konnte also mit den Prädiktoren in Modell 08 um 11% reduziert werden. Betrachtet man jedoch die AIC- und BIC-Werte (vgl. Kap. B5), so sind diese in Modell 08 zwar niedriger als die Werte in Modell 01, sie sind aber wieder deutlich höher als die AIC/BIC-Werte von Modell 07. Das Modell 08 ist also eindeutig weniger gut den Daten angepasst als das Modell 07, was sicherlich eine Folge des Informationsverlustes ist, der eintritt, wenn PA dichotomisiert wird und als PA2 nur noch einen von zwei möglichen Werten aufweisen kann.
- 4) Die L1-Varianz ist mit  $p=0,000$  hochsignifikant (vgl. Tab. 8.2). Das Schätzmodell könnte also noch um zusätzliche Prädiktoren erweitert werden, um freie Varianzanteile zu binden.
- 5) Die Varianz der LEZU-Intercepts und die Varianz der PA2-Slopes sind hochsignifikant (vgl. Tab. 8.2). Beide Größen (LEZU-Intercepts und PA2-Slopes) variieren demnach sehr deutlich zwischen den verschiedenen Ländern/Staaten. Für den SEX2-Slope gilt dies nicht. Er variiert zwischen den Ländern/Staaten in nicht-signifikanter Weise ( $p=0,183$ ).
- 6) Die Kovarianzen zwischen allen Residuen sind nicht signifikant. Es macht also keinen Sinn, diese Werte inhaltlich interpretieren zu wollen.
- 7) Viele Modellannahmen, die für eine ML-Schätzung des Modells 08 gelten sollten (z.B. Abwesenheit von Multikollinearität u.a.), wurden hier aus Zeitgründen nicht untersucht, sollten allerdings üblicherweise bei einer ML-Modellschätzung getestet werden (vgl. dazu Kap. A2).

Abb. 8.3: Graphische Skizze des Gleichungssystems von Modell 08 mit Schätzergebnaten



Tab. 8.1: Ergebnisse der MLA-Schätzung von Modell 08: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) Mean (LEZU)	3,995	0,000
L2-Slope1, ( $\gamma_{10}$ ) Mean Slope2(PA2)	0,761	0,000
L2-Slope2, ( $\gamma_{20}$ ) Mean Slope1(SEX2)	0,138	0,000

Tab. 8.2: Ergebnisse der MLA-Schätzung von Modell 08: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	1,084	0,000
L2-Varianz, LEZU, Var( $u_{0j}$ )	0,216	0,000
L2-Varianz, Slope1(PA2) Var( $u_{1j}$ )	0,034	0,005
L2-Varianz, Slope2(SEX2) Var( $u_{2j}$ )	0,012	0,183

Tab. 8.3: Ergebnisse der MLA-Schätzung von Modell 08: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}$ , $u_{1j}$ )	0,003	0,826
Cov( $u_{0j}$ , $u_{2j}$ )	-0,028	0,100
Cov( $u_{1j}$ , $u_{2j}$ )	0,005	0,611

Tab. 8.4: Mplus-Inputfile für Modell 08

```

TITLE:
  Modell "MLA 08"
  mit zwei L1-Prädiktoren: "pa2" (mit Mediansplit !!) und "sex2" (0/1)
  mit keinem L2-Prädiktor
  mit random intercept und mit zwei random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa2   ! pa2 ist X-Var: positiver Affekt (Mediansplit)
  sex2; ! 0/1 = m/w

WITHIN ARE pa2 sex2 ; ! pa2 und sex2 werden als L1-Prädiktoren definiert

CLUSTER=nation;

DEFINE:
  IF(sex EQ 1) THEN sex2=0;
  IF(sex EQ 2) THEN sex2=1;

  ! hier kommt der Mediansplit von PA:
  IF(pa LE 5.67 )THEN pa2=0;
  IF(pa GT 5.67) THEN pa2=1;

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
                          ! wenn random slopes geschätzt werden sollen
  MODEL: %WITHIN%
    Beta1j | lezu ON pa2; ! Spezifikation der L1-Regression mit Random Slope
              ! Beta1j ist Name für Random Slope
    Beta2j | lezu ON sex2 ; ! Spezifikation der L1-Regression mit Random Slope
              ! Beta2j ist Name für Random Slope

  %BETWEEN%
    lezu WITH Beta1j; ! Kovarianz zwischen Random Intercept und random slope
    lezu WITH Beta2j; ! Kovarianz zwischen Random Intercept und random slope
    Beta1j WITH Beta2j; ! Kovarianz zw. X1-random slope und X2-random slope

OUTPUT: sampstat;

```

Tab. 8.5: Mplus-Ausgabe von Modellschätzung 08 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				10
Loglikelihood				
H0 Value				-10197.339
H0 Scaling Correction Factor for MLR				1.2070
Information Criteria				
Akaike (AIC)				20414.678
Bayesian (BIC)				20483.113
Sample-Size Adjusted BIC				20451.336
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	1.084	0.034	31.700	0.000
Between Level				
LEZU WITH				
BETA1J	0.003	0.015	0.220	0.826
BETA2J	-0.028	0.017	-1.643	0.100
BETA1J WITH				
BETA2J	0.005	0.009	0.509	0.611
Means				
LEZU	3.995	0.077	52.082	0.000
BETA1J	0.761	0.041	18.465	0.000
BETA2J	0.138	0.033	4.198	0.000
Variances				
LEZU	0.216	0.052	4.157	0.000
BETA1J	0.034	0.012	2.812	0.005
BETA2J	0.012	0.009	1.332	0.183



Modell 09:

Zwei-Ebenen-Modell,

Random-Intercepts- / Fixed+Random-Slopes-Modell

mit L1-Prädiktor "PA2" (dichotomisiert) und Random PA2-Slopes

mit L1-Prädiktor "SEX2" (dichotomisiert, 0/1=m/w) und Fixed SEX2-Slopes

ohne L2-Prädiktoren

L1-Gleichung  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (9.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + u_{0j}$  (9.2a)

$\beta_{1j} = \gamma_{10} + u_{1j}$  (9.2b)

$\beta_{2j} = \gamma_{20}$  (9.2c)

“fixed part”:  $\gamma_{00}, \gamma_{10}, \gamma_{20}$

“random part”:  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$

“covariance”:  $\text{Cov}(u_{0j}, u_{1j})$

(Das Modell 09 ist identisch mit Modell 08, allerdings jetzt mit Fixed Slopes für den SEX2-Effekt.)

Abb. 9.1: Graphische Skizze des Gleichungssystems von Modell 09

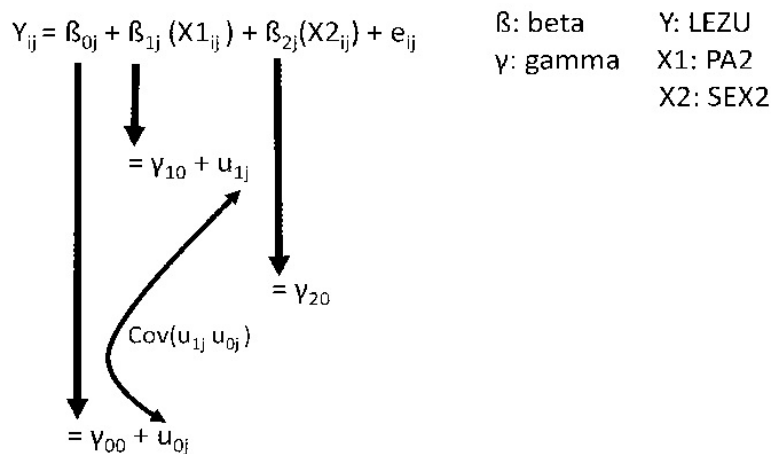


Abb. 9.2a: Stilisierte Darstellung der Modell-Logik von Modell 09 (Version 1)

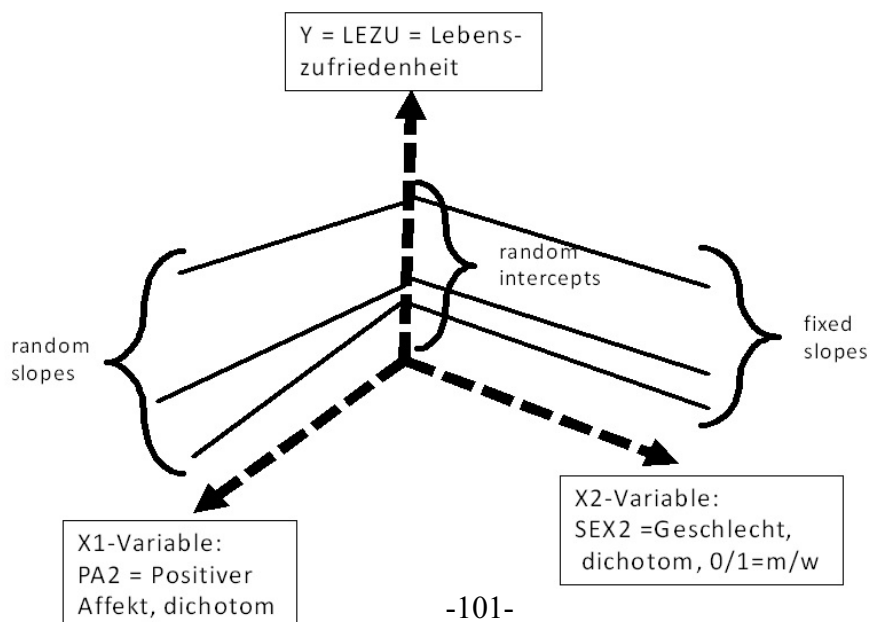
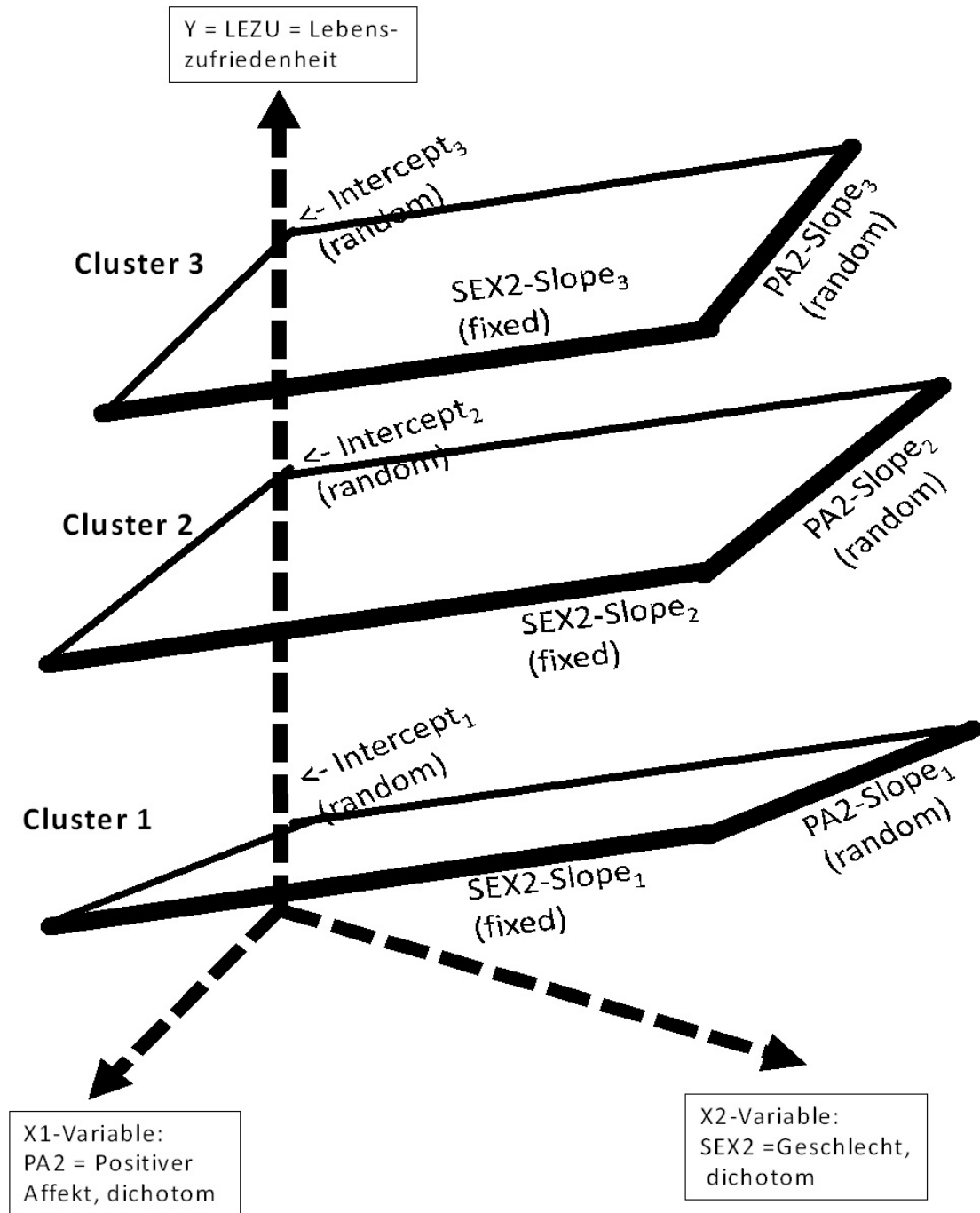


Abb. 9.2b: Stilisierte Darstellung der Modell-Logik von Modell 09 (Version 2)



Zur Logik von Modell 09:

Das Modell 09 unterscheidet sich von Modell 08 nur dadurch, dass nunmehr die SEX2-Slopes nicht wie im Modell 08 von Cluster zu Cluster variieren, sondern über alle Cluster hinweg konstant bleiben (fixed slopes). Demgegenüber variieren die PA2-Slopes nach wie vor (d.h. wie auch schon in Modell 08) länderspezifisch (= random slopes). Vergleicht man das Modell 09 mit Modell 07, so fällt auf, dass dort die SEX2-Slopes nicht fix sind. Zudem ist der L1-Prädiktor "PA2" in Modell 09 mittels Mediansplit dichotomisiert, während er in Modell 07 noch metrisch (zentriert) ist. Mit der Modellschätzung soll ermittelt werden:

- Wie hoch wird die durchschnittliche LEZU mit Modell 09 über alle Länder hinweg geschätzt?
- Wie stark sind die Effekte von PA2 und SEX2, welcher Effekt ist stärker?
- Sollte der SEX2-Effekt als random effect (wie in Modell 08) oder als fixed effect (wie in Modell 09) spezifiziert und geschätzt werden?
- Haben die "random parts" des Modells genügend freie Varianz, um die Erklärungsleistung des Modells noch durch zusätzliche Parameter zu vergrößern?

Die Abbildungen 9.2a und 9.2b verdeutlichen die Logik von Modell 09. Wir konzentrieren uns hier auf die Abbildung 9.2b und verweisen auf unsere umfassende Erläuterung der Graphik im Textabschnitt "Die Abbildungen 3.2 und 3.3 verdeutlichen ...". Im Vergleich zu Abbildung 3.3 sind in Abbildung 9.2b nunmehr zwei Veränderungen vorgenommen worden: a) die Variable PA ist nicht zentriert sondern dichotomisiert, b) der PA2-Slope ist nicht als "fixed" sondern als "random" definiert worden.

Zur Interpretation des Gleichungssystems von Modell 09:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (9.1)$$

Y abh. L1-Variable "LEZU=Lebenszufriedenheit"

$\beta_{0j}$  L1-Intercept in der Lineargleichung zur Schätzung von LEZU mittels der Prädiktoren "SEX2 (dichotom)" und "PA2=Positiver Affekt (dichotom)". Das Intercept informiert über den Referenzwert zur Beurteilung der geschätzten Slopes. Ist PA2=0 (=unterdurchschnittl. PA-Wert wg. Mediansplit) und ist gleichzeitig SEX2=0 (=m), so hat LEZU einen Wert von 3,988 (=Referenzwert).

$\beta_{1j}$  partieller PA2-Slope in der Lineargleichung zur Schätzung von LEZU mittels Prädiktor "PA2 (dichotomisiert)"

X1 unabh. L1-Prädiktor "PA2 (dichotom)"

$\beta_{2j}$  partieller SEX2-Slope in der Lineargleichung zur Schätzung von Y mittels Prädiktor "SEX2 (dichotom)"

X2 unabh. L1-Prädiktor "SEX2=0/1= männl./weibl. (dichotom)"

$e_{ij}$  Differenz zwischen geschätztem und empirischem Y-Wert (L1-Residualvariable)

Var( $e_{ij}$ ) Varianz der Residualvariable  $e_{ij}$  (s.o.) auf der L1-Ebene. Wird auch als Within-Varianz bezeichnet (Varianz innerhalb der Cluster/Staaten/Länder). Die Varianz betrifft die Streuung der Abweichungen der individuellen LEZU-Werte vom jeweiligen länderspezifischen LEZU-Mittelwert.

Mplus: Within Level / Variances LEZU = 1,086 (p=0,000)

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (9.2a)$$

$\gamma_{00}$  Durchschnitt (Grand Mean) von allen länderspezifischen Intercepts. Mittleres Intercept über alle Länder hinweg.

Mplus: Between Level / Means LEZU = 3,988 (p=0,000)

$u_{0j}$  Abweichungen der länderspezifischen Intercepts vom Durchschnitt aller Intercepts (L2-Residualvariable)

Var( $u_{0j}$ ) Varianz von  $u_{0j}$

Mplus: Between Level / Variances LEZU = 0,191 (p=0,000)

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (9.2b)$$

$\gamma_{10}$  Durchschnitt von allen länderspezifischen partiellen PA2-Slopes. Mittlerer Slope über alle Länder hinweg

Mplus: Between Level / Means BETA1J = 0,763 (p=0,000)

$u_{1j}$  Differenzen der länderspezifischen partiellen PA2-Slopes vom Durchschnittswert aller partiellen PA2-Slopes. Die Residualvariable macht deutlich, dass der geschätzte PA2-Effekt nicht konstant ist, sondern von Land zu Land variiert.

Var( $u_{1j}$ ) Varianz von  $u_{1j}$

Mplus: Between Level / Variances BETA1J = 0,035 (p=0,006)

$$\beta_{2j} = \gamma_{20} \text{ (Schätzung eines länderbezogen konstanten SEX2-Slopes = fixed slope)} \quad (9.2c)$$

$\gamma_{20}$  Die Effektstärke, mit der X2 (=SEX2) die abh. Variable Y (=LEZU) beeinflusst ( $\beta_{2j}$ ), ist zwar im Prinzip länderspezifisch ( $\beta_{2j}$ ) und kann wie ein unstandardisierter partieller Regressionskoeffizient interpretiert werden. Wir gehen aber davon aus, dass die Effektstärke nicht variiert, sondern in jedem Land gleich ist und eine feste L2-Größe " $\gamma_{20}$ " annimmt (fixed slope-Annahme).

Mplus: Within Level / LEZU ON SEX2: 0,142 (p=0,000)

Cov( $u_{0j}$ ,  $u_{1j}$ ) Kovarianz zwischen länderspezifischen Intercepts und länderspezifischen partiellen PA2-Slopes

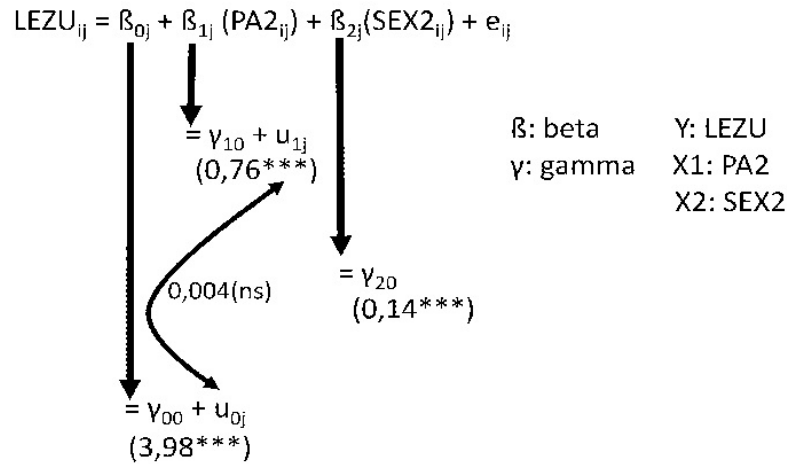
Mplus: Between Level / LEZU WITH BETA1J = 0,004 (p=0,758)

Zur Interpretation der Schätzung von Modell 09

(vgl. dazu auch die folgende Abbildung 9.3 sowie die Tabellen 9.1 bis 9.3)

- 1) Ein durchschnittlicher, über alle Länder hinweg gültiger LEZU-Wert lässt sich mit der Spezifikation des Modells 09 nicht berechnen. Der geschätzte Intercept-Wert (3,988) informiert über die Ausprägung von LEZU für männliche Personen ( $SEX2=0=m$ ), die einen unterdurchschnittlichen PA2-Wert aufweisen ( $PA2=0$ , Mediansplit). Von daher kann  $\gamma_{00}$  als länderübergreifender LEZU-Referenzwert interpretiert werden. So steigt der LEZU-Wert im Durchschnitt um 0,142 Einheiten an (SEX2-Slope), wenn nicht die Gruppe der männlichen Befragten, sondern die Gruppe der weiblichen Befragten betrachtet wird. Auch steigt der LEZU-Wert für Befragte mit überdurchschnittlichem PA2-Wert (im Unterschied zu Befragten mit unterdurchschnittlichem LEZU-Wert) um 0,763 Einheiten an.
- 2) Da SEX2 und PA2 dichotomisierte Prädiktoren sind, lassen sich die Slopes von beiden Prädiktoren direkt miteinander vergleichen (0,142 vs. 0,763). Von daher ist der Effekt von PA2 ca. fünfmal so stark wie der Effekt von SEX2.
- 3) Der Slope von SEX2 wurde in Modell 09 als fixer Parameter geschätzt, von dem angenommen wurde, dass er in jedem der analysierten Länder den gleichen Wert aufweist. Der BIC-Wert unterstützt diese Annahme, weil er von Modell 08 zu Modell 09 deutlich abfällt (von 20483,113 auf 20462,977), allerdings bleibt der AIC-Wert fast konstant (M8: 20414,678, M9: 20415,072) (zu AIC/BIC vgl. auch Kap. B5).
- 4) Die Varianzschätzungen aller Residuen sind hochsignifikant, sodass in die Modellschätzung (unter rein statistischen Gesichtspunkten) noch weitere “erklärende” Prädiktoren aufgenommen werden können. Auch kann somit gerechtfertigterweise angenommen werden, dass die Intercepts und die PA2-Slopes länderspezifisch variieren.
- 5) Viele Modellannahmen, die für eine ML-Schätzung des Modells 09 gelten sollten (z.B. Abwesenheit von Multikollinearität u.a.) wurden hier aus Zeitgründen nicht untersucht, sollten allerdings üblicherweise bei einer ML-Modellschätzung getestet werden (vgl. dazu Kap. A2).

Abb. 9.3: Graphische Skizze des Gleichungssystems von Modell 09 mit Schätzergebnaten



Tab. 9.1: Ergebnisse der MLA-Schätzung von Modell 09: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) Mean (LEZU)	3,988	0,000
L2-Slope1, ( $\gamma_{10}$ ) Mean Slope1(PA2)	0,763	0,000
L1-Slope (fixed), ( $\gamma_{20}$ ) LEZU ON SEX2	0,142	0,000

Tab. 9.2: Ergebnisse der MLA-Schätzung von Modell 09: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	1,086	0,000
L2-Varianz, LEZU, Var( $u_{0j}$ )	0,191	0,000
L2-Varianz, Slope1(PA2) Var( $u_{1j}$ )	0,035	0,006

Tab. 9.3: Ergebnisse der MLA-Schätzung von Modell 09: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}$ , $u_{1j}$ )	0,004	0,758

Tab. 9.4: Mplus-Inputfile für Modell 09

```

TITLE:
  Modell "MLA 09"
  mit zwei L1-Prädiktoren: "pa2" (mit Mediansplit) und "sex2" (0/1)
  mit keinem L2-Prädiktor
  mit random intercept und mit random slope für PA2
  mit fixed slope für SEX2

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa2  ! PA2 ist X-Var: positiver Affekt (Mediansplit)
  sex2; ! 0/1 = m/w

WITHIN = pa2 sex2 ; ! PA2 und SEX2 werden als L1-Prädiktoren definiert
BETWEEN = ;
CLUSTER = nation;

DEFINE:
  IF(sex EQ 1) THEN sex2=0;
  IF(sex EQ 2) THEN sex2=1;

  IF(pa LE 5.67 )THEN pa2=0; ! hier ist der Mediansplit von PA:
  IF(pa GT 5.67) THEN pa2=1;

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
                          ! wenn random slopes geschätzt werden sollen

MODEL: %WITHIN%
  Beta1j | lezu ON pa2; ! Spezifikation der L1-Regression mit Random Slope
                ! Beta1j ist Name für Random Slope

  lezu ON sex2 ; !      ! geschätzt wird fixed slope auf L1

%BETWEEN%
  lezu WITH beta1j;

OUTPUT: sampstat;

```

Tab. 9.5: Mplus-Ausgabe von Modellschätzung 09 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				7
Loglikelihood				
H0 Value				-10200.536
H0 Scaling Correction Factor for MLR				1.3262
Information Criteria				
Akaike (AIC)				20415.072
Bayesian (BIC)				20462.977
Sample-Size Adjusted BIC				20440.732
(n* = (n + 2) / 24)				
MODEL RESULTS				
	Estimate	S.E.	Two-Tailed Est./S.E.	P-Value
Within Level				
LEZU ON				
SEX2	0.142	0.033	4.249	0.000
Residual Variances				
LEZU	1.086	0.034	31.763	0.000
Between Level				
LEZU WITH				
BETA1J	0.004	0.014	0.308	0.758
Means				
LEZU	3.988	0.078	50.888	0.000
BETA1J	0.763	0.041	18.580	0.000
Variances				
LEZU	0.191	0.044	4.306	0.000
BETA1J	0.035	0.013	2.748	0.006



Modell 10:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell,

mit einem L1-Prädiktor "PA" (zentriert),

mit einem L2-Prädiktor "GDP2" (dichotomisiert) zur Schätzung von Intercepts  
 ("intercepts-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + e_{ij}$  (10.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z1_j + u_{0j}$  (10.2a)

$\beta_{1j} = \gamma_{10} + u_{1j}$  (10.2b)

"fixed part":  $\gamma_{00}, \gamma_{01}, \gamma_{10}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$

"covariance":  $\text{Cov}(u_{0j}, u_{1j})$

Abb. 10.1: Graphische Skizze des Gleichungssystems von Modell 10

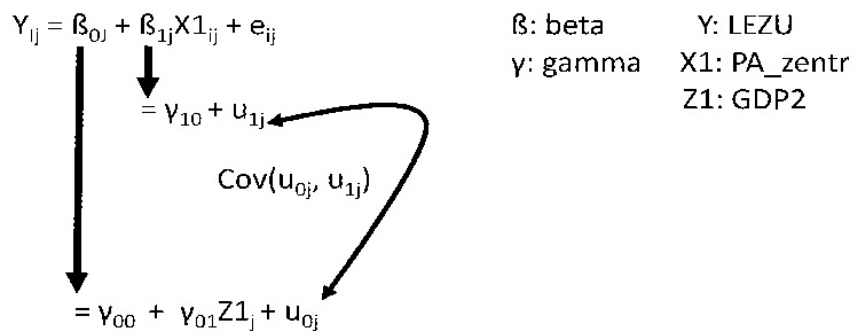
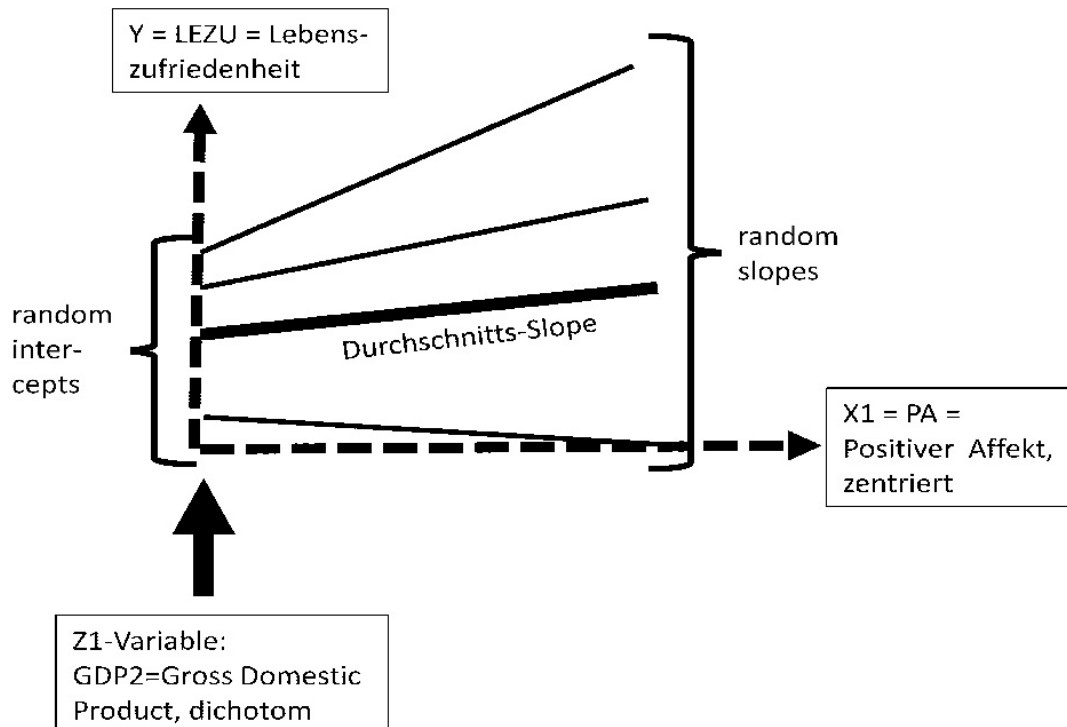


Abb. 10.2: Stilisierte Darstellung der Modell-Logik von Modell 10



Zur Logik von Modell 10:

Das Modell 10 enthält sowohl einen L1-Prädiktor "X1=PA\_zentriert" als auch einen L2-Prädiktor "Z1=GDP2\_dichotomisiert". Der Prädiktor "Z1=GDP2" bezeichnet das nach Mediansplit dichotomisierte Bruttonationaleinkommen auf Staatsebene und variiert von Staat zu Staat.

Im Modell beeinflusst Z1=GDP2 den Achsenabschnitt (Intercept) der Gleichung zur Schätzung der abhängigen L1-Variablen Y=LEZU. Mit der Modellschätzung soll ermittelt werden:

- Auf welchem durchschnittlichen Niveau befindet sich die individuelle LEZU?
- Wie stark wird das Niveau von LEZU von dem Prädiktor GDP2 beeinflusst?
- Wie stark beeinflusst PA die individuelle LEZU?
- Variieren die Prädiktor-Effekte in signifikanter Weise zwischen den Ländern?

Die Abbildung 10.2 verdeutlicht die Logik von Modell 10. Das Koordinatensystem von Modell 10 ist identisch mit demjenigen von Modell 6. Zusätzlich weist Modell 10 noch den dichotomen L2-Prädiktor "Z1/GDP2" auf, der die länderspezifisch variierenden Intercepts beeinflussen soll (fett gezeichneter Pfeil von Z1 auf die Intercepts).

Zur Interpretation des Gleichungssystems von Modell 10:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + e_{ij} \quad (10.1)$$

Y abh. Variable LEZU auf der Mikroebene L1

$\beta_{0j}$  Achsenabschnitt/Intercept der Schätzung von Y (s.u.)

$\beta_{1j}$  Slope des Effekts von X1 auf Y auf der Mikroebene (s.u.)

X1 unabh. Variable PA auf der Mikroebene L1 (zentriert mit Grand Mean)

$e_{ij}$  Residualvariable, Diff. zwischen beob. LEZU und geschätzter LEZU auf Personenebene

Var( $e_{ij}$ ) Varianz der Residualvariable  $e_{ij}$  (s.o.)

Mplus: Within Level / Residual Variances LEZU = 0,995 (p=0,000)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z1_j + u_{0j} \quad (10.2a)$$

$\gamma_{00}$  Intercept der Gleichung zur Schätzung aller länderspezifischen Intercepts (ist erwarteter globaler LEZU-Wert bei X1=0 und Z1=0)

Mplus: Between Level / Intercepts LEZU = 4,170 (p=0,000)

$\gamma_{01}$  geschätzter Anstieg im Intercept, wenn sich der L2-Prädiktor GDP2 (0/1) von 0 auf 1 erhöht (bei  $\gamma_{01} = 0$ : kein Unterschied zwischen den GDP2-Gruppen; bei  $\gamma_{01} > 0$ : Intercept in Gruppe 1 ist größer als Intercept in Gruppe 0; bei  $\gamma_{01} < 0$ : Intercept in Gruppe 0 ist größer als in Gruppe 1)

Mplus: Between Level / LEZU On GDP2 = 0,494 (p=0,000)

Z1 L2-Prädiktor Z1=GDP2 (dichotom: 0/1)

$u_{0j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspez. Intercepts  $\beta_{0j}$  und den mittels Gl. 10.2a (unter Einsatz von Prädiktor Z1=GDP2) geschätzten Intercepts

Var( $u_{0j}$ ) Varianz von  $u_{0j}$  (s.o.)

Mplus: Between Level / Residual Variances LEZU = 0,116 (p=0,000)

$$\beta_{ij} = \gamma_{10} + u_{ij} \quad (10.2b)$$

Gleichung zur Schätzung eines PA-Slope-Durchschnittswertes auf Länderebene

$\gamma_{10}$  Durchschnittswert von PA-Slope auf Länderebene

Mplus: Between Level / Means BETA1J = 0,349 (p = 0,000)

$u_{ij}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Slopes  $\beta_{ij}$  und durchschnittlichem Slope  $\gamma_{10}$  (auf Länderebene)

Var( $u_{ij}$ ) Varianz der L2-Residualvariable  $u_{ij}$  (s.o.)

Mplus: Between Level / Variances BETA1J = 0,006 (p=0,000)

Cov( $u_{0j}$ ,  $u_{1j}$ ) Kovarianz zwischen länderspezifischen Intercepts und PA-Slopes

Mplus: Between Level / LEZU WITH BETA1J = 0,006 (p=0,241)

Zur Interpretation der Schätzung von Modell 10:

(vgl. dazu auch die folgende Abbildung 10.3 sowie die Tabellen 10.1 bis 10.3)

- 1) Das geschätzte, individuelle Intercept der Gleichung zur Schätzung von LEZU beträgt dann, wenn alle Prädiktoren einen Wert von 0 aufweisen: 4,17 (LEZU-Skala: 1-7). Der L1-Prädiktor "PA" ist zentriert und ist deshalb dann null, wenn PA seinen Mittelwert annimmt. Der L2-Prädiktor "GDP2" ist dichotomisiert und nimmt einen Wert von null an, wenn das GDP eines Landes kleiner oder gleich des GDP-Medianwertes ist (Median=0,74; GDP-Skala: 0,36- 0,94)
- 2) Das geschätzte mittlere Niveau von LEZU (=4,17) erhöht sich um 0,49 Punkte, wenn die Befragten in Ländern mit überdurchschnittlichen GDPs wohnen.
- 3) Der Effekt von PA auf LEZU erhöht sich um jeweils 0,34 Punkte, wenn der PA-Wert der Befragten um eine Skaleneinheit ansteigt.
- 4) Die Varianzen der geschätzten Parameter sind alle hoch signifikant (vgl. Tab. 10.2), d.h. die Schätzwerte variieren bedeutsam zwischen den untersuchten Ländern.
- 5) Zur Inspektion der Signifikanz aller geschätzten Modellgrößen vgl. die folgende Abbildung. Die Markierung der Schätzwerte mit \*\*\* bedeutet, dass diese als hochsignifikant (p=0,000) anzusehen sind. Bei einer Irrtumswahrscheinlichkeit > 0,05 und damit bei einer nicht vorhandenen Signifikanz wird der entsprechende Schätzwert mit (ns) gekennzeichnet.

Abb. 10.3: Graphische Skizze des Gleichungssystems von Modell 10 mit Schätzergebnaten

$$LEZU_{ij} = \beta_{0j} + \beta_{1j}(PA\_zentr_{ij}) + e_{ij}$$

$\downarrow$   
 $= \gamma_{10} + u_{1j}$   
 (0,34\*\*\*)  
 $\leftarrow$  0,006(ns)

$$= \gamma_{00} + \gamma_{01}(GDP2)_j + u_{0j}$$

(4,17\*\*\*) (0,49\*\*\*)  
 (0,006)

$\beta$ : beta

$\gamma$ : gamma

Y: LEZU

X1: PA\_zentr

Z1: GDP2

Tab. 10.1: Ergebnisse der MLA-Schätzung von Modell 10: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	4,170	0,000
L2-Slope1, ( $\gamma_{10}$ ) Mean Slope1(PA_zentr)	0,349	0,000
L2-Slope, ( $\gamma_{01}$ ) LEZU ON GDP2	0,494	0,000

Tab. 10.2: Ergebnisse der MLA-Schätzung von Modell 10: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	0,995	0,000
L2-Residualvarianz, LEZU, Var( $u_{0j}$ )	0,116	0,000
L2-Varianz, Slope1(PA_zentr) Var( $u_{1j}$ )	0,006	0,000

Tab. 10.3: Ergebnisse der MLA-Schätzung von Modell 10: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}$ , $u_{1j}$ )	0,006	0,241

Tab. 10.4: Mplus-Inputfile für Modell 10

```

TITLE:
  Modell "MLA 10"
  mit 1 L1-Prädiktor: "pa-zentriert"
  mit 1 L2-Prädiktor: gdp2 (für L1-Intercept)
  mit random intercepts und mit random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu ! lezu ist die abh. Y-Variable
  pa !
  gdp2; !

WITHIN ARE pa ; ! pa wird als L1-Prädiktor definiert

BETWEEN ARE gdp2; ! definiert gdp2 als L2-Prädiktor (dichotomisiert)

CLUSTER=nation;

DEFINE:
  ! hier kommt der Mediansplit:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74) THEN gdp2=1;

  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert
  ! keine Zentrierung bei Dummies

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
  ! wenn random slopes geschätzt werden sollen

MODEL: %WITHIN%
  Beta1j | lezu ON pa ; ! Spezifikation der L1-Regression mit Random Slope

  %BETWEEN%
  lezu ON gdp2; ! Schätzung des Intercepts durch L2-Prädiktor
  lezu WITH Beta1j; ! Residual-Kovarianz zwischen Random Intercept
  ! und random slope

OUTPUT:
  sampstat;

```

Tab. 10.5: Mplus-Ausgabe von Modellschätzung 10 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				7
Loglikelihood				
H0 Value				-9891.882
H0 Scaling Correction Factor for MLR				1.2859
Information Criteria				
Akaike (AIC)				19797.763
Bayesian (BIC)				19845.667
Sample-Size Adjusted BIC				19823.423
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	0.995	0.034	29.241	0.000
Between Level				
LEZU ON				
GDP2	0.494	0.112	4.424	0.000
LEZU WITH				
BETA1J	0.006	0.005	1.173	0.241
Means				
BETA1J	0.349	0.016	22.041	0.000
Intercepts				
LEZU	4.170	0.084	49.856	0.000
Variances				
BETA1J	0.006	0.001	4.323	0.000
Residual Variances				
LEZU	0.116	0.024	4.859	0.000

Modell 11:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell (mit cross-level-Effekt),

mit L1-Prädiktor "PA"(zentriert),

mit L2-Prädiktor "GDP2"(dichotomisiert, Mediansplit),

zur Schätzung von Intercepts und Slopes

("intercepts-and-slopes-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + e_{ij}$  (11.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z1_j + u_{0j}$  (11.2a)

$\beta_{1j} = \gamma_{10} + \gamma_{11}Z1_j + u_{1j}$  (11.2b)

"fixed part":  $\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$

"covariance"  $\text{Cov}(u_{0j}, u_{1j})$

(Im Unterschied zu Modell 10 wird hier der Effekt von Z1 (=GDP2) sowohl auf die Intercepts als auch auf die PA-Slopes (das ist neu) geschätzt.)

Abb. 11.1: Graphische Skizze des Gleichungssystems von Modell 11

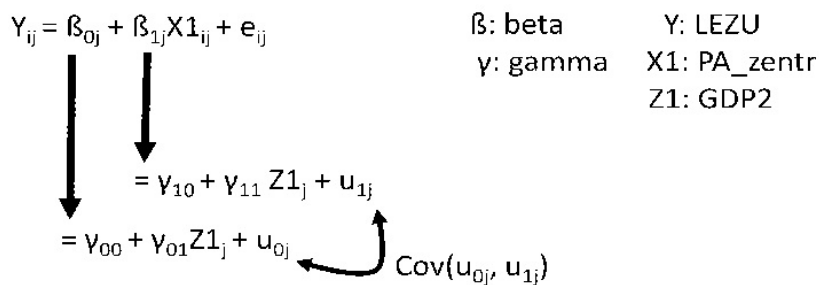
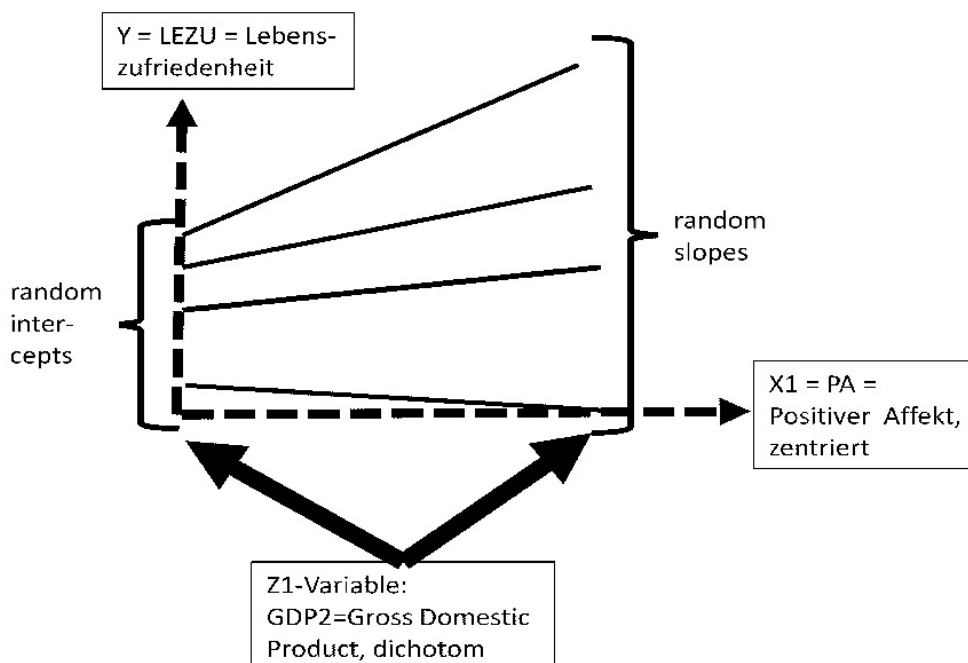


Abb.11.2: Stilisierte Darstellung der Modell-Logik von Modell 11



Zur Logik von Modell 11:

Das Modell 11 enthält sowohl einen L1-Prädiktor "PA\_zentriert" als auch einen L2-Prädiktor "GDP2\_dichotomisiert". Im Gleichungssystem werden diese beiden Prädiktoren als "X1" und "Z1" bezeichnet (Gl. 11.2a/11.2b). In unserem Beispiel ist Z1=GDP2 (nach Mediansplit dichotomisiertes Bruttosozialprodukt auf Staatsebene). Die L2-Variable GDP2 variiert von Staat zu Staat. Die Variable bleibt aber für jeden Staat konstant.

Im Modell beeinflusst Z1=GDP2 den Achsenabschnitt (Intercept) und den PA\_zentr-Slope der Gleichung zur Schätzung der abhängigen L1-Variablen Y=LEZU. Mit der Modellschätzung soll ermittelt werden:

- Auf welchem durchschnittlichen Niveau befindet sich die individuelle LEZU?
- Wie stark wird das Niveau von LEZU von den Prädiktoren PA\_zentr und GDP2 beeinflusst?
- Wie stark wird der Slope (der Steigungskoeffizient) von PA\_zentr in der Gleichung zur Schätzung von LEZU von dem L2-Prädiktor GDP2 beeinflusst (d.h. wie stark wirkt der cross-level-Effekt)?

Die Abbildung 11.2 verdeutlicht die Logik von Modell 11. Das Koordinatensystem von Modell 11 (gestrichelte Linien) ist identisch mit demjenigen von Modell 6. Zusätzlich weist Modell 11 aber noch den dichotomen L2-Prädiktor "Z1/GDP2" auf, der die länderspezifisch variierenden Intercepts sowie die ebenfalls länderspezifisch variierenden Slopes des Effekts von X1/PA\_zentr auf Y (Linien mit positiver Steigung im Koordinatensystem) beeinflusst (fett gezeichneter Pfeil von Z1 auf die Intercepts und auf die Slopes).

Zur Interpretation des Gleichungssystems von Modell 11:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + e_{ij} \quad (11.1)$$

L1-Gleichung zur Schätzung von Y=LEZU durch X1=PA\_zentr

Y abh. Variable LEZU auf der Mikroebene L1

$\beta_{0j}$  Achsenabschnitt/Intercept der Schätzung von Y (s.u.)

$\beta_{1j}$  Slope des Effekts von X1 auf Y auf der Mikroebene (s.u.)

X1 unabh. Variable PA\_zentr auf der Mikroebene L1 (zentriert mit Grand Mean)

$e_{ij}$  Residualvariable: Differenz zwischen beobachteter LEZU und geschätzter LEZU auf L1-Personenebene.

Var( $e_{ij}$ ) Varianz der Residualvariable  $e_{ij}$  (s.o.)

Mplus: Within Level / Residual Variances LEZU = 0,995 (p=0,000)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z1_j + u_{0j} \quad (11.2a)$$

L2-Gleichung zur Schätzung von  $\beta_{0j}$  durch Z1=GDP2

$\gamma_{00}$  Intercept der Gleichung zur Schätzung aller länderspezifischen Intercepts (ist erwarteter globaler LEZU-Wert bei  $X_1=0$  und  $Z=0$ )

Mplus: Between Level / Intercepts LEZU = 4,151 (p=0,000)

$\gamma_{01}$  geschätzter Anstieg im Intercept, wenn sich der L2-Prädiktor GDP2 (0/1) von 0 auf 1 erhöht (bei  $\gamma_{01} = 0$ : kein Unterschied zwischen den GDP2-Gruppen; bei  $\gamma_{01} > 0$ : Intercept in Gruppe 1 ist größer als Intercept in Gruppe 0; bei  $\gamma_{01} < 0$ : Intercept in Gruppe 0 ist größer als in Gruppe 1)



Mplus: Between Level / LEZU On GDP2 = 0,533 (p=0,000)

Z1 L2-Prädiktor Z=GDP2 (dichotom: 0/1)

$u_{0j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Intercepts  $\beta_{0j}$  und den mittels Gl. 11.2a (unter Einsatz von Prädiktor Z=GDP2) geschätzten Intercepts, sodass es sich also um Differenzen handelt, die nicht mit L2-Prädiktoren (Z-Variablen) erklärt werden können.

Var( $u_{0j}$ ) Varianz von  $u_{0j}$  (s.o.)

Mplus: Between Level / Residual Variances LEZU = 0,116 (p=0,000)

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z1_j + u_{1j} \quad (11.2b)$$

Gleichung zur Schätzung eines GDP2-Effekts auf den PA\_zentr-Slope

$\gamma_{10}$  erwarteter PA\_zentr-Slope, wenn GDP2=0 ist

Mplus: Between Level / Means BETA1J = 0,319 (p = 0,000)

$\gamma_{11}$  Interaktionseffekt zwischen L1-Prädiktor (PA\_zentr) und L2-Prädiktor (GDP2) (auch: cross-level-Effekt); geschätzter Effekt des GDP2-Steigungskoeffizienten auf den Slope von PA\_zentr, hier als Differenz zwischen GDP=0 und GDP=1; geschätzter Anstieg des PA\_zentr-Slopes, wenn der L2-Prädiktor GDP2 (0/1) von 0 auf 1 erhöht wird (bei  $\gamma_{11} = 0$ : kein Unterschied zwischen den GDP2-Gruppen; bei  $\gamma_{11} > 0$ : Slope in Gruppe 1 ist größer als Slope in Gruppe 0; bei  $\gamma_{11} < 0$ : Slope in Gruppe 0 ist größer als in Gruppe 1)

Mplus: Between Level / BETA1J ON GDP2= 0,061 (p = 0,044)

$u_{1j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Slopes  $\beta_{1j}$  und den mittels Gl. 11.2b (unter Einsatz von Prädiktor Z=GDP2) geschätzten Slopes, sodass es sich also um Differenzen handelt, die nicht mit L2-Prädiktoren (Z-Variablen) erklärt werden können

Var( $u_{1j}$ ) Varianz der Residualvariable  $u_{1j}$  (s.o.)

Mplus: Between Level / Residual Variances BETA1J = 0,006 (p=0,000)

Cov( $u_{0j}$ ,  $u_{1j}$ ) Kovarianz zwischen den zwei L2-Residualvariablen (s.o.)

Mplus: Between Level / LEZU WITH BETA1J = 0,006 (p=0,229)

### Zur Interpretation der Schätzung von Modell 11

(vgl. dazu auch die folgende Abbildung 11.3 sowie die Tabellen 11.1 bis 11.3)

- 1) Das Intercept der Gleichung zur Schätzung von LEZU bzw. der Durchschnittswert von LEZU beträgt dann, wenn alle Prädiktoren einen Wert von 0 aufweisen: 4,15 (LEZU-Skala: 1 - 7). Der L1-Prädiktor PA ist zentriert und ist null, wenn PA seinen Mittelwert annimmt. Der L2-Prädiktor GDP2 ist dichotomisiert und nimmt einen Wert von null an, wenn das GDP eines Landes kleiner oder gleich dem GDP-Medianwert ist (Median=0,74; GDP-Skala: 0,36 - 0,94).
- 2) Das geschätzte mittlere Niveau von LEZU (=4,15, s.o.) erhöht sich um 0,53 Punkte, wenn die Befragten aus Ländern mit überdurchschnittlichen GDPs kommen.
- 3) Die LEZU-Steigerungsrate von 0,32 Punkten bei Erhöhung von PA\_zentr um eine Skaleneinheit vergrößert sich noch einmal um 0,06 Punkte auf insgesamt 0,38 Punkte, wenn GDP2 von 0 auf 1 ansteigt. Es gibt also zwei verschiedene PA\_zentr-Effekte bei Schätzung eines statistisch erwarteten LEZU-Wertes: einen PA\_zentr-Effekt für die Gruppe der Befragten aus Ländern mit unterdurchschnittlichen GDPs (Slope=0,32) und einen PA\_zentr-Effekt für die Gruppe der Befragten aus Ländern mit überdurchschnittlichen GDPs (Slope=0,38). Der GDP2-Effekt verursacht also einen cross-level-Effekt, der mit einem  $p=0,044$  nur schwach signifikant ist.
- 4) Auch die AIC- und BIC-Werte von Modell 10 und Modell 11 unterscheiden sich kaum und signalisieren somit, dass der Fit der beiden Modelle sehr ähnlich ist (vgl. Kap. B5).
- 5) Zur Inspektion der Signifikanz aller geschätzten Modellgrößen vgl. die folgende Abbildung. Die Markierung der Schätzwerte mit \*\*\* bedeutet, dass diese als hochsignifikant ( $p=0,000$ ) anzusehen sind. Eine schwache Signifikanz ist mit \* ausgezeichnet (z.B. 0,061\*). Bei einer Irrtumswahrscheinlichkeit  $> 0,05$  und damit bei einer nicht vorhandenen Signifikanz wird der entsprechende Schätzwert mit (ns) gekennzeichnet.

Abb. 11.2: Graphische Skizze des Gleichungssystems von Modell 11 mit Schätzergebnaten

$$LEZU_{ij} = \beta_{0j} + \beta_{1j}(PA\_zentr_{ij}) + e_{ij}$$

$\beta$ : beta      Y: LEZU  
 $\gamma$ : gamma    X1: PA\_zentr  
 Z1: GDP2

$$= \gamma_{10} + \gamma_{11}(GDP2_j) + u_{1j}$$

(0,31\*\*\*) (0,06\*)

$$= \gamma_{00} + \gamma_{01}(GDP2_j) + u_{0j}$$

(4,15\*\*\*) (0,53\*\*\*)

0,006(ns)

Tab. 11.1: Ergebnisse der MLA-Schätzung von Modell 11: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	4,151	0,000
L2-Intercept, ( $\gamma_{10}$ ) Slope1(PA_zentr)	0,319	0,000
L2-Slope, ( $\gamma_{01}$ ) LEZU ON GDP2	0,533	0,000
L2-Slope1, ( $\gamma_{11}$ ) Slope1 ON GDP2	0,061	0,044

Tab. 11.2: Ergebnisse der MLA-Schätzung von Modell 11: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	0,995	0,000
L2-Residualvarianz, LEZU, Var( $u_{0j}$ )	0,116	0,000
L2-Residualvarianz, Slope1, Var( $u_{1j}$ )	0,006	0,000

Tab. 11.3: Ergebnisse der MLA-Schätzung von Modell 11: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
Cov( $u_{0j}$ , $u_{1j}$ )	0,006	0,229

Tab. 11.4: Mplus-Inputfile für Modell 11

```

TITLE:
  Modell "MLA 11"
  mit L1-Prädiktor: "pa_zentriert"
  mit L2-Prädiktor: gdp2 zur Schätzung von Intercepts + Slopes
                    dafür mit random intercepts & mit random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
            KLA_F  LEZU  PA  NA  BAL  IC
            GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
            PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa   ! unabh. Var. auf L1 (X1)
  gdp2; ! unabh. Var. auf L2 (Z1)

WITHIN ARE pa ; ! pa wird als L1-Prädiktor definiert

BETWEEN ARE gdp2; ! definiert gdp2 als L2-Prädiktor (dichotomisiert)

CLUSTER=nation;

DEFINE:
  ! hier kommt der Mediansplit:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74 )THEN gdp2=1;

  CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert
                        ! Achtung: keine Zentrierung bei Dummies

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
                          ! wenn random slopes geschätzt werden sollen

MODEL: %WITHIN%
  Beta1j | lezu ON pa ; ! Spezifikation der L1-Regression mit Random Slope
                    ! Beta1j ist Name für Random Slope

  %BETWEEN%
  lezu ON gdp2;    ! Schätzung der Intercepts durch L2-Prädiktor
  Beta1j ON gdp2; ! Schätzung des Slope durch L2-Prädiktor

  lezu WITH Beta1j; ! Residual-Kovarianz zwischen Random Intercept und random slope

OUTPUT:
  sampstat;

```

Tab. 11.5: Mplus-Ausgabe von Modellschätzung 11 (ausgewählte Bereiche):

MODEL FIT INFORMATION				
Number of Free Parameters				8
Loglikelihood				
H0 Value				-9889.966
H0 Scaling Correction Factor for MLR				1.2654
Information Criteria				
Akaike (AIC)				19795.931
Bayesian (BIC)				19850.679
Sample-Size Adjusted BIC				19825.257
(n* = (n + 2) / 24)				
MODEL RESULTS				
			T wo-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	0.995	0.034	29.255	0.000
Between Level				
BETA1J	ON			
GDP2	0.061	0.030	2.016	0.044
LEZU	ON			
GDP2	0.533	0.112	4.773	0.000
LEZU	WITH			
BETA1J	0.006	0.005	1.203	0.229
Intercepts				
LEZU	4.151	0.084	49.625	0.000
BETA1J	0.319	0.023	13.941	0.000
Residual Variances				
LEZU	0.116	0.024	4.862	0.000
BETA1J	0.006	0.001	3.843	0.000

Modell 12:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell (mit cross-level-Effekten),

mit L1-Prädiktor "PA2" (dichotomisiert, Mediansplit),

mit L1-Prädiktor "SEX2" (dichotomisiert),

mit L2-Prädiktor "GDP2"(dichotomisiert, Mediansplit)

zur Schätzung von Intercepts und Slopes  
("intercepts-and-slopes-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (12.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (12.2a)

$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$  (12.2b)

$\beta_{2j} = \gamma_{20} + \gamma_{21}Z_j + u_{2j}$  (12.2c)

"fixed part":  $\gamma_{00}, \gamma_{10}, \gamma_{20}, \gamma_{01}, \gamma_{11}, \gamma_{21}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j}), \text{Var}(u_{2j})$

"covariance":  $\text{Cov}(u_{0j}, u_{1j}), \text{Cov}(u_{0j}, u_{2j}), \text{Cov}(u_{1j}, u_{2j})$

(Modell 12 enthält im Unterschied zu Modell 11 den zusätzlichen X-Prädiktor "SEX2". Auch sind im Unterschied zu Modell 11 nunmehr alle Prädiktoren dichotomisiert (PA2, SEX2, GDP2). Bei PA2 und GDP2 erfolgte die Dichotomisierung in Form eines Mediansplits. Der Prädiktor SEX2 wurde in Form von 0/1=m/w dichotomisiert.)

Abb 12.1: Graphische Skizze des Gleichungssystems von Modell 12

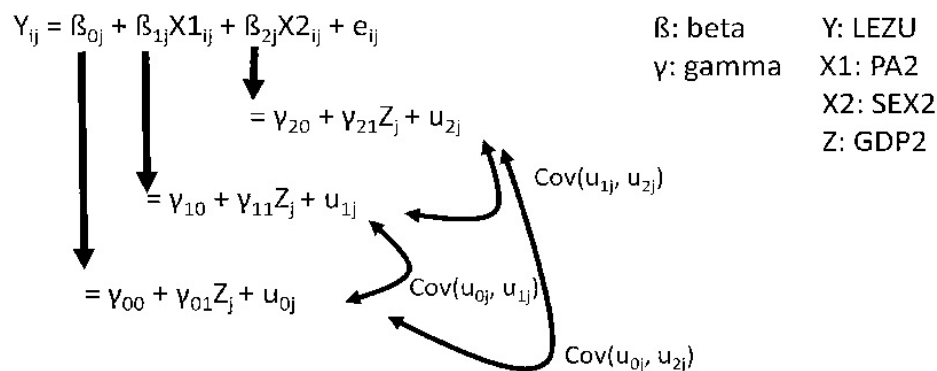
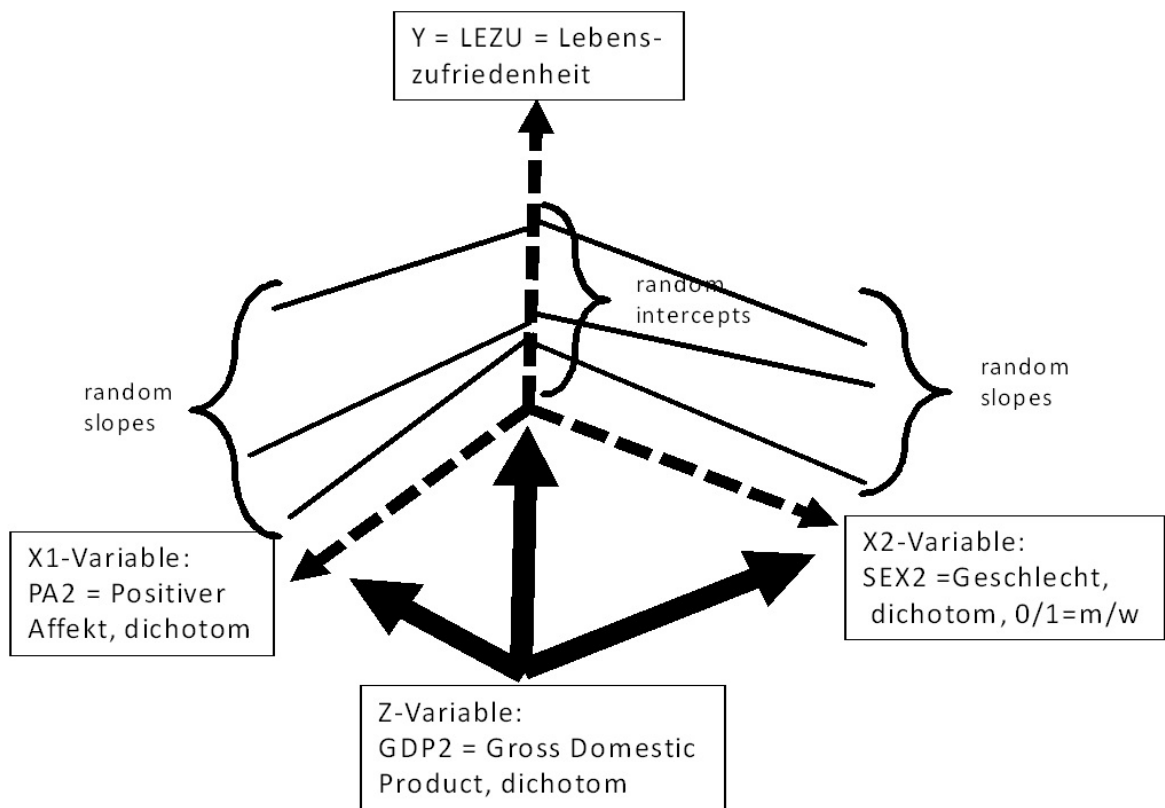


Abb. 12.2: Stilisierte Darstellung der Modell-Logik von Modell 12

Zur Logik von Modell 12:

Modell 12 enthält im Unterschied zu Modell 11 zwei X-Prädiktoren, die nunmehr auch beide dichotomisiert sind: der L1-Prädiktor PA2 (Mediansplit) sowie der L1-Prädiktor SEX2 (0/1 = m/w). Da auch der L2-Prädiktor GDP2 (gross national product) dichotomisiert ist (Mediansplit), können alle Prädiktoren hinsichtlich ihrer Effektstärken zur Beeinflussung von LEZU direkt miteinander verglichen werden. Dies gilt auch für alle cross-level-Effekte, nach denen der L2-Prädiktor Z/GDP2 auf die Slopes von X1/PA2 und X2/SEX2 sowie auf das LEZU-Intercept einwirkt.

Die Abbildung 12.2 verdeutlicht die Logik von Modell 12. Das dreidimensionale Koordinatensystem von Modell 12 (gestrichelte Linien) betrifft die L1-Variablen: Y/LEZU, X1/PA2 und X2/SEX2. Im Koordinatensystem sind die random X1- und X2-Slopes (durchgezogene Linien mit unterschiedlichen Steigungen) sowie die dazugehörigen Intercepts eingezeichnet. Zusätzlich weist Modell 12 auch noch den dichotomen L2-Prädiktor "Z1/GDP2" auf, der die länderspezifisch variierenden Intercepts sowie die ebenfalls länderspezifisch variierenden X1- und X2-Slopes beeinflusst (fett gezeichnete Pfeile von Z1 auf die Intercepts und auf die Slopes).

Zur Interpretation des Gleichungssystems:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij} \quad (12.1)$$

Ist eine L1-Gleichung zur Schätzung von  $Y=LEZU$  durch  $X1=PA2$  und  $X2=SEX2$ .

$Y$  abh. Variable  $LEZU$  auf der Mikroebene L1

$\beta_{0j}$  Achsenabschnitt/Intercept der Schätzung von  $Y$  (s.u.)

$\beta_{1j}$  Slope des Effekts von  $X1$  auf  $Y$  auf der Mikroebene (s.u.)

$\beta_{2j}$  Slope des Effekts von  $X2$  auf  $Y$  auf der Mikroebene (s.u.)

$X1$  unabh. Variable  $PA2$  auf der Mikroebene L1 (dichotomisiert nach Mediansplit)

$X2$  unabh. Variable  $SEX2$  auf der Mikroebene L1 (0/1 = m/w)

$e_{ij}$  Residualvariable: Differenz zwischen beobachteter  $LEZU$  und geschätzter  $LEZU$  auf L1

$Var(e_{ij})$  Varianz der Residualvariable  $e_{ij}$  (s.o.)

Mplus: Within Level / Residual Variances  $LEZU = 1,083$  ( $p=0,000$ )

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (12.2a)$$

Ist eine L2-Gleichung zur Schätzung von  $\beta_{0j}$  durch  $Z=GDP2$ .

$\gamma_{00}$  Intercept der Gleichung zur Schätzung aller länderspezifischen Intercepts (ist erwarteter globaler  $LEZU$ -Wert bei  $X1, X2, Z=0$ )

Mplus: Between Level / Intercepts  $LEZU = 3,768$  ( $p=0,000$ )

$\gamma_{01}$  geschätzter Anstieg des Intercepts, wenn sich der L2-Prädiktor  $GDP2$  (0/1) von 0 auf 1 erhöht (bei  $\gamma_{01} = 0$ : kein Unterschied zwischen den  $GDP2$ -Gruppen; bei  $\gamma_{01} > 0$ : Intercept in Gruppe 1 ist größer als Intercept in Gruppe 0; bei  $\gamma_{01} < 0$ : Intercept in Gruppe 0 ist größer als in Gruppe 1)

Mplus: Between Level /  $LEZU$  On  $GDP2 = 0,452$  ( $p=0,000$ )

$Z$  L2-Prädiktor  $Z=GDP2$  (dichotom: 0/1)

$u_{0j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Intercepts  $\beta_{0j}$  und den mittels Gl. 12.2a (unter Einsatz von Prädiktor  $Z=GDP2$ ) geschätzten Intercepts, sodass es sich also um Differenzen handelt, die nicht mit L2-Prädiktoren ( $Z$ -Variablen) erklärt werden können

$Var(u_{0j})$  Varianz von  $u_{0j}$  (s.o.)

Mplus: Between Level / Residual Variances  $LEZU = 0,165$  ( $p=0,000$ )

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (12.2b)$$

Ist eine Gleichung zur Schätzung eines  $GDP2$ -Effekts auf den  $PA2$ -Slope.

$\gamma_{10}$  erwarteter  $PA2$ -Slope, wenn  $GDP2=0$  ist

Mplus: Between Level / Intercepts  $BETA1J = 0,714$  ( $p = 0,000$ )



$\gamma_{11}$  Interaktionseffekt zwischen PA2-Effekt (L1) und GDP2-Effekt (L2) (auch: cross-level-Effekt); geschätzter Effekt des GDP2-Steigungskoeffizienten auf den Slope von PA2, hier als Differenz zwischen GDP2=0 und GDP2=1; geschätzter Anstieg des PA2-Slopes, wenn sich der L2-Prädiktor GDP2 (0/1) von 0 auf 1 erhöht (bei  $\gamma_{11} = 0$ : kein Unterschied zwischen den GDP2-Gruppen; bei  $\gamma_{11} > 0$ : Slope in Gruppe 1 ist größer als Slope in Gruppe 0; bei  $\gamma_{11} < 0$ : Slope in Gruppe 0 ist größer als in Gruppe 1)

Mplus: Between Level / BETA1J ON GDP2= 0,097 (p = 0,212)

$u_{1j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Slopes  $\beta_{1j}$  und den mittels Gl. 12.2b (unter Einsatz von Prädiktor Z=GDP2) geschätzten Slopes, sodass es sich also um Differenzen handelt, die nicht mit dem L2-Prädiktor (Z-Variable) erklärt werden können

Var( $u_{1j}$ ) Varianz der Residualvariable  $u_{1j}$  (s.o.)

Mplus: Between Level / Residual Variances BETA1J = 0,033 (p=0,007)

$$\beta_{2j} = \gamma_{20} + \gamma_{21}Z_j + u_{2j} \quad (12.2c)$$

Ist eine Gleichung zur Schätzung eines GDP2-Effekts auf den SEX2-Slope.

$\gamma_{20}$  erwarteter SEX2-Slope, wenn GDP2=0 ist

Mplus: Between Level / Intercepts BETA2J = 0,137 (p = 0,009)

$\gamma_{21}$  Interaktionseffekt zwischen SEX2-Effekt (L1) und GDP2-Effekt (L2) (auch: cross-level-Effekt); geschätzter Effekt des GDP2-Steigungskoeffizienten auf den Slope von SEX2, hier als Differenz zwischen GDP2=0 und GDP2=1; geschätzter Anstieg des PA2-Slopes, wenn sich der L2-Prädiktor GDP2 (0/1) von 0 auf 1 erhöht (bei  $\gamma_{11} = 0$ : kein Unterschied zwischen den GDP2-Gruppen; bei  $\gamma_{11} > 0$ : Slope in Gruppe 1 ist größer als Slope in Gruppe 0; bei  $\gamma_{11} < 0$ : Slope in Gruppe 0 ist größer als in Gruppe 1)

Mplus: Between Level / BETA2J ON GDP2= 0,005 (p = 0,941)

$u_{2j}$  L2-Residualvariable: Differenzen zwischen geschätzten, länderspezifischen Slopes  $\beta_{2j}$  und den mittels Gl. 12.2c (unter Einsatz von Prädiktor Z=GDP2) geschätzten Slopes, sodass es sich also um Differenzen handelt, die nicht mit dem L2-Prädiktor (Z-Variable) erklärt werden können

Var( $u_{2j}$ ) Varianz der Residualvariablen  $u_{2j}$  (s.o.)

Mplus: Between Level / Residual Variances BETA2J = 0,012 (p=0,175)

Cov( $u_{0j}$ ,  $u_{1j}$ ) Kovarianz zwischen den zwei L2-Residualvariablen (s.o.)

Mplus: Between Level / LEZU WITH BETA1J = -0,008 (p=0,524)

Cov( $u_{0j}$ ,  $u_{2j}$ ) Kovarianz zwischen den zwei L2-Residualvariablen (s.o.)

Mplus: Between Level / LEZU WITH BETA2J = -0,029 (p=0,070)

Cov( $u_{1j}$ ,  $u_{2j}$ ) Kovarianz zwischen den zwei L2-Residualvariablen (s.o.)

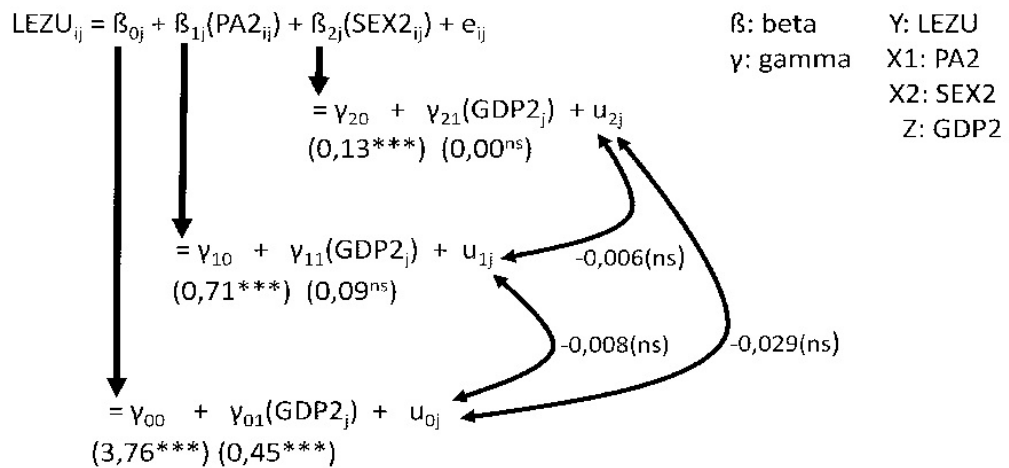
Mplus: Between Level / BETA1J WITH BETA2J = 0,006 (p=0,504)

Zur Interpretation der Schätzung von Modell 12:

(vgl. dazu auch die folgende Abbildung 12.3 sowie die Tabellen 12.1/12.2/12.3 )

- 1) Das Intercept der Gleichung zur Schätzung von LEZU bzw. der Durchschnittswert von LEZU beträgt dann, wenn alle Prädiktoren einen Wert von 0 aufweisen: 3,76 (LEZU-Skala: 1-7). Dies ist dann der Fall, wenn der dichotome L1-Prädiktor PA2 (0/1) einen unterdurchschnittl. Wert aufweist, wenn der dichotome L1-Prädiktor SEX2 (0/1) die Ausprägung "männl." annimmt und wenn der dichotome L2-Prädiktor GDP2 (0/1) einen unterdurchschnittl. Wert hat.
- 2) Das geschätzte mittlere Niveau von LEZU (=3,76, s.o.) erhöht sich signifikant um 0,45 Punkte, wenn die Befragten aus Ländern mit überdurchschnittlichen GDPs kommen.
- 3) Wenn PA2 von 0 auf 1 ansteigt (von unterdurchschnittl. auf überdurchschnittl.), vergrößert sich LEZU signifikant um 0,71 Pkte. Diese Punktzahl erhöht sich noch einmal um 0,09 Pkte, wenn GDP2 von 0 auf 1 ansteigt (von unterdurchschnittl. auf überdurchschnittl.). Dies wäre ein cross-level-Effekt. Allerdings ist der Zuschlag von 0,09 Punkten nicht signifikant und deshalb gibt es bezügl. des PA2-Prädikt. keinen von GDP2 ausgehenden cross-level-Effekt.
- 4) Wenn SEX2 von 0 auf 1 ansteigt (von männl. auf weibl.), vergrößert sich LEZU signifikant um 0,13 Punkte. Frauen haben also einen signifikant höheren LEZU-Wert als Männer. Der Anstieg könnte sich noch einmal um 0,005 Punkte erhöhen, wenn Länder mit überdurchschnittlichem GDP (d.h. mit GDP2=1) betrachtet werden. Allerdings ist dieser zusätzliche Anstieg nicht signifikant und deshalb gibt es auch bezüglich des SEX2-Prädiktors keinen von GDP2 ausgehenden cross-level-Effekt.
- 5) Es gibt in Modell 12 keine signifikanten, von GDP2 (Level 2) ausgehenden und auf PA2 bzw. SEX2 (Level 1) ausgerichteten cross-level-Effekte.
- 6) Der Effekt von PA2 auf LEZU (0,71) ist hochsignifikant und bei GDP2=0 um ca. fünfmal stärker als der hochsignifikante Effekt von SEX2 auf LEZU (0,13). Der hochsignifikante Effekt von GDP2 auf LEZU (0,45) ist ca. ein Drittel schwächer als der Effekt von PA2. Es gilt also: Effekt(PA2) > Effekt(GDP2) > Effekt(SEX2). Dieser Vergleich ist möglich, weil alle L1- und L2-Prädiktoren dichotomisiert wurden.
- 7) Viele Modellannahmen, die für eine ML-Schätzung des Modells 12 gelten sollten (z.B. Abwesenheit von Multikollinearität) wurden hier aus Zeitgründen nicht untersucht, sollten allerdings üblicherweise bei einer ML-Modellschätzung getestet werden (vgl. dazu Kap. A2).
- 8) Zur Inspektion der Signifikanz aller geschätzten Modellgrößen vgl. die folgende Abbildung. Die Markierung der Schätzwerte mit \*\*\* bedeutet, dass diese als hochsignifikant ( $p=0,000$ ) anzusehen sind. Eine schwache Signifikanz ist mit \* ausgezeichnet (z.B. 0,061\*). Bei einer Irrtumswahrsch. > 0,05 und damit bei einer nicht vorhandenen Signifikanz wird der entsprechende Schätzwert mit (ns) gekennzeichnet.

Abb. 12.3: Graphische Skizze des Gleichungssystems von Modell 12 mit Schätzergebnaten



Tab. 12.1: Ergebnisse der MLA-Schätzung von Modell 12: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	3,768	0,000
L2-Intercept, ( $\gamma_{10}$ ) Slope1 (PA2)	0,714	0,000
L2-Intercept, ( $\gamma_{20}$ ) Slope2 (SEX2)	0,137	0,000
L2-Slope, ( $\gamma_{01}$ ) LEZU ON GDP2	0,452	0,00
L2-Slope, ( $\gamma_{21}$ ) Slope2 ON GDP2	0,005	0,941
L2-Slope, ( $\gamma_{11}$ ) Slope1 ON GDP2	0,097	0,212

Tab. 12.2: Ergebnisse der MLA-Schätzung von Modell 12: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, $Var(e_{ij})$	1,083	0,000
L2-Residualvarianz, LEZU, $Var(u_{0j})$	0,165	0,000
L2-Residualvarianz, Slope1(PA2) $Var(u_{1j})$	0,033	0,007
L2-Residualvarianz, Slope2(SEX2) $Var(u_{2j})$	0,012	0,175

Tab. 12.3: Ergebnisse der MLA-Schätzung von Modell 12: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
$\text{Cov}(u_{0j}, u_{1j})$	-0,008	0,524
$\text{Cov}(u_{0j}, u_{2j})$	-0,029	0,070
$\text{Cov}(u_{1j}, u_{2j})$	0,006	0,504

Tab. 12.4: Mplus-Inputfile für Modell 12

```

TITLE:
  Modell "MLA 12"
  mit L1-Prädiktor: "PA2" (dichotom, Mediansplit)
  mit L1-Prädiktor: "SEX2" (dichotom)
  mit L2-Prädiktor: "gdp2" (dichotom, Mediansplit)
    (zur Schätzung von Intercepts + Slopes)
  mit random intercepts, mit zwei random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu ! lezu ist die abh. Y-Variable
  pa2 !
  sex2 !
  gdp2; !

WITHIN ARE pa2 sex2; ! pa2 + sex2 werden als L1-Prädiktoren definiert

BETWEEN ARE gdp2; ! definiert gdp2 als L2-Prädiktor (dichotomisiert)

CLUSTER=nation;

DEFINE:
  ! statt 1/2 nunmehr 0/1-Dichotomisierung
  IF(sex EQ 1) THEN sex2=0; ! 0=männl, 1=weibl
  IF(sex EQ 2) THEN sex2=1;

  ! hier kommt Mediansplit:
  IF(pa LE 5.67 )THEN pa2=0;
  IF(pa GT 5.67) THEN pa2=1;

  ! hier kommt der Mediansplit:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74) THEN gdp2=1;

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
    ! wenn random slopes geschätzt werden sollen

```

FORTSETZUNG ->

Fortsetzung von Tabelle 12.4:

MODEL: %WITHIN%

Beta1j Beta2j | lezu ON pa2 sex2 ; ! Spezifikation der L1-Regression  
! mit zwei Random Slopes,  
! Beta1j ist frei gewählter Name für Random Slope

%BETWEEN%

lezu beta1j beta2j ON gdp2; ! Schätzung der Intercepts und der Slopes  
! durch L2-Prädiktor gdp2

lezu WITH Beta1j; ! Residual-Kovarianz zwischen  
! Random Intercept und random slope1

lezu WITH Beta2j;  
Beta2j WITH Beta1j;

OUTPUT:

sampstat;

Tab. 12.5: Mplus-Ausgabe von Modellschätzung 12 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				13
Loglikelihood				
H0 Value				-10189.621
H0 Scaling Correction Factor for MLR				1.1494
Information Criteria				
Akaike (AIC)				20405.241
Bayesian (BIC)				20494.206
Sample-Size Adjusted BIC				20452.895
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	1.083	0.034	31.749	0.000
Between Level				
BETA1J	ON			
GDP2	0.097	0.078	1.248	0.212
BETA2J	ON			
GDP2	0.005	0.071	0.074	0.941
LEZU	ON			
GDP2	0.452	0.141	3.216	0.001
LEZU	WITH			
BETA1J	-0.008	0.013	-0.637	0.524
BETA2J	-0.029	0.016	-1.814	0.070
BETA2J	WITH			
BETA1J	0.006	0.009	0.668	0.504
Intercepts				
LEZU	3.768	0.099	38.090	0.000
BETA1J	0.714	0.067	10.699	0.000
BETA2J	0.137	0.052	2.616	0.009
Residual Variances				
LEZU	0.165	0.038	4.370	0.000
BETA1J	0.033	0.012	2.681	0.007
BETA2J	0.012	0.009	1.356	0.175

Modell 13:

Zwei-Ebenen-Modell,

Random-Intercepts-/Random-Slopes-Modell (mit cross-level-Effekt),

mit L1-Prädiktor "PA2" (dichotomisiert, Mediansplit),

mit L1-Prädiktor "SEX2" (dichotomisiert),

mit L2-Prädiktor "GDP2"(dichotomisiert, Mediansplit)

zur Schätzung von Intercepts und Slopes

("intercepts-and-slopes-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (13.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (13.2a)

$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$  (13.2b)

$\beta_{2j} = \gamma_{20} + u_{2j}$  (13.2.c)

"fixed part":  $\gamma_{00}, \gamma_{10}, \gamma_{20}, \gamma_{01}, \gamma_{11}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j}), \text{Var}(u_{2j})$

"covariance":  $\text{Cov}(u_{0j}, u_{1j}), \text{Cov}(u_{0j}, u_{2j}), \text{Cov}(u_{1j}, u_{2j})$

(Das Modell 13 unterscheidet sich nur dadurch von Modell 12, dass die Variation der Random Slopes von X2=SEX2 nicht auf Z=GDP2 zurückgeführt wird und es somit keinen Prädiktor für X2 gibt.)

Abb. 13.1 Graphische Skizze des Gleichungssystems von Modell 13

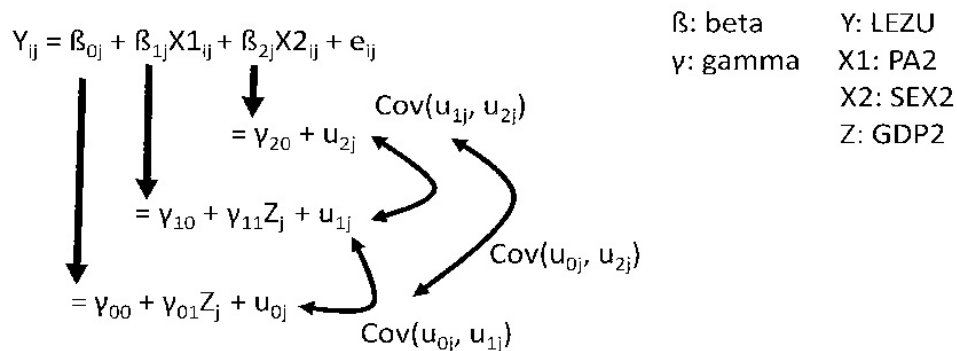
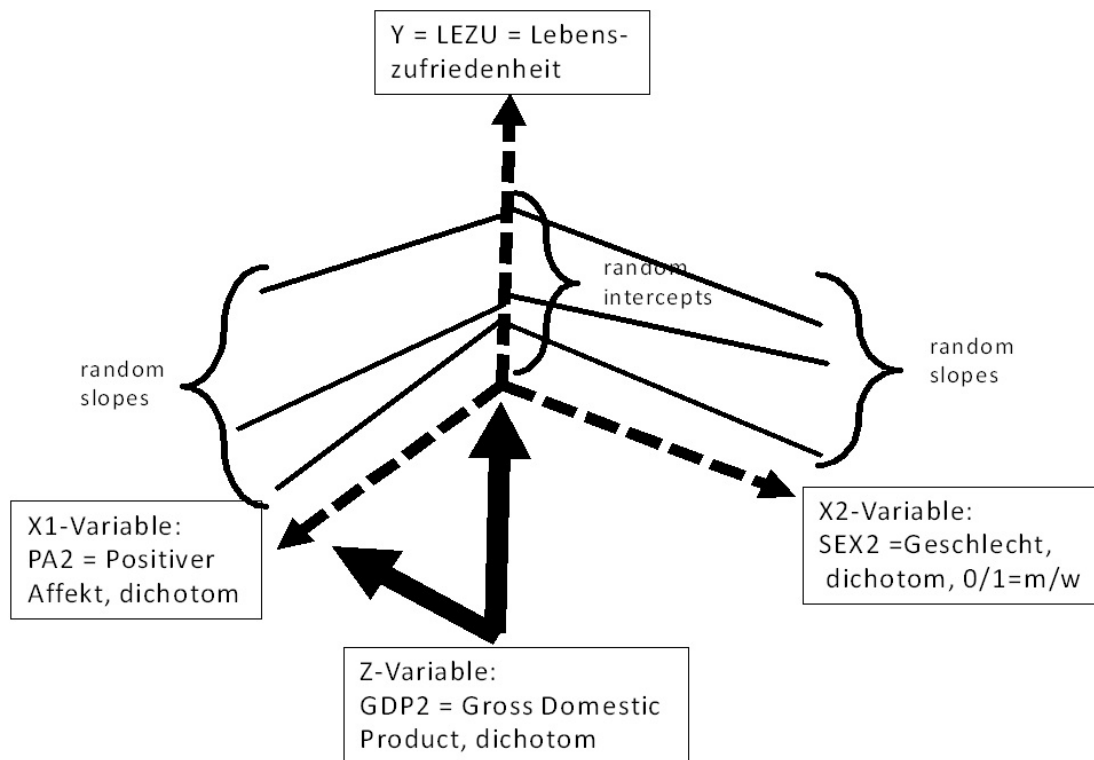




Abb. 13.2: Stilisierte Darstellung der Modell-Logik von Modell 13



### Zur Logik von Modell 13:

In Modell 13 wird angenommen, dass der SEX2-Effekt unabhängig von GDP2-Einflüssen ist, was auch dem Schätzergebnis von Modell 12 entspricht, in dem der Effekt von GDP2 auf die Slopes von SEX2 als nicht-signifikant geschätzt wurde (vgl. Abbildung 12.3). Es wird allerdings auch angenommen, dass der SEX2-Effekt zwischen den einzelnen Ländern variiert und deshalb die SEX2-Slopes als "random" angesehen werden müssen. Es werden deshalb die ersten drei Gleichungen von Modell 12 (12.1, 12.2a, 12.2b) auch für die Spezifikation von Modell 13 übernommen (13.1, 13.2a, 13.2b). Allein die Gleichung 12.2c wird durch die neue Gleichung 13.2c (s.o.) ersetzt.

Im Folgenden wird deshalb allein die Gleichung 13.2c erläutert. Zum Verständnis der Gleichungen 13.1 bis 13.2b können die diesbezüglichen Ausführungen bei Modell 12 benutzt werden. Die einzelnen Parameter-Schätzwerte, die sehr ähnlich den Schätzwerten von Modell 12 sind, können der folgenden Abbildung 13.3 entnommen werden.

Die Abbildung 13.2 verdeutlicht die Logik von Modell 13. Die Darstellung basiert wiederum auf einem dreidimensionalen Koordinatensystem mit den drei Variablen Y/LEZU, X1/PA2 und X2/SEX2 (gestrichelte Linien). Das Modell 13 unterscheidet sich dadurch von Modell 12, dass die Variation der Random Slopes von X2/SEX2 (durchgezogene Linien mit unterschiedlichen Steigungen) nicht auf Z/GDP2 zurückgeführt wird und es somit keinen Prädiktor für X2 gibt.

Zur Interpretation des Gleichungssystems von Modell 13:

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (13.2c)$$

Ist Gleichung zur Schätzung eines mittleren SEX2-Slopes.

$\gamma_{20}$  Mittelwert aller länderspezifischen SEX2-Slopes

Mplus: Between Level / Means / BETA2J = 0,139 (p = 0,000)

$u_{2j}$  L2-Residualvariable: Differenzen zwischen den geschätzten, länderspezifischen Slopes und dem mittleren, durchschnittlichen Slope

Var( $u_{2j}$ ) Varianz der Residualvariablen  $u_{2j}$  (s.o.), gleichzeitig auch die Varianz der SEX2-Slopes

Mplus: Between Level / Variances BETA2J = 0,012 (p=0,176)

(für weitere Erläuterungen zu den Modellgleichungen 13.1 bis 13.2b vgl. die Ausführungen zu den Gleichungen 12.1 bis 12.2b)

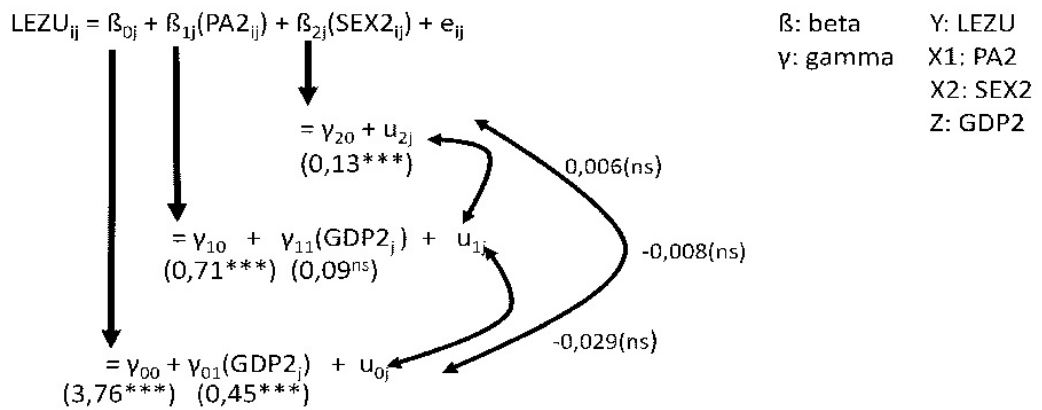
Zur Interpretation der Schätzung von Modell 13

(vgl. dazu auch die folgende Abbildung 13.3 sowie die Tabellen 13.1 bis 13.3)

- 01) Das Intercept der Gleichung zur Schätzung von LEZU bzw. der Durchschnittswert von LEZU beträgt dann, wenn alle Prädiktoren einen Wert von 0 aufweisen: 3,76 (LEZU-Skala: 1 - 7). Dies ist dann der Fall, wenn der dichotomisierte L1-Prädiktor PA2 (0/1) einen unterdurchschnittlichen Wert aufweist, wenn der dichotomisierte L1-Prädiktor SEX2 (0/1) die Ausprägung "männlich" annimmt und wenn der dichotomisierte L2-Prädiktor GDP2 (0/1) einen unterdurchschnittlichen Wert betrifft.
- 02) Das geschätzte mittlere Niveau von LEZU (=3,76, s.o.) erhöht sich signifikant um 0,45 Punkte, wenn die Befragten aus Ländern mit überdurchschnittlichen GDPs kommen.
- 03) Wenn PA2 von 0 auf 1 ansteigt (von unterdurchschnittlich auf überdurchschnittlich), vergrößert sich LEZU signifikant um 0,71 Punkte. Diese Punktzahl erhöht sich noch einmal um 0,09 Punkte, wenn GDP2 von 0 auf 1 ansteigt (von unterdurchschnittlich auf überdurchschnittlich). Dies wäre ein cross-level-Effekt. Allerdings ist der Zuschlag von 0,09 Punkten nicht signifikant und deshalb gibt es bezüglich des PA2-Prädiktors keinen von GDP2 ausgehenden cross-level-Effekt.
- 04) Wenn SEX2 von 0 auf 1 ansteigt (von männlich auf weiblich), vergrößert sich LEZU signifikant um 0,13 Punkte. Frauen haben also einen signifikant höheren LEZU-Wert als Männer.
- 05) Es gibt in Modell 13 keinen signifikanten, von GDP2 (Level 2) ausgehenden und auf PA2 ausgerichteten cross-level-Effekt.

- 06) Der Effekt von PA2 auf LEZU (0,71) ist hochsignifikant und bei  $GDP2=0$  um ca. fünfmal stärker als der hochsignifikante Effekt von SEX2 auf LEZU (0,13). Der hochsignifikante Effekt von GDP2 auf LEZU (0,45) ist um ca. ein Drittel schwächer als der Effekt von PA2 (0,71). Es gilt also:  $\text{Effekt}(PA2) > \text{Effekt}(GDP2) > \text{Effekt}(SEX2)$ . Dieser Vergleich ist möglich, weil alle L1- und L2-Prädiktoren dichotomisiert wurden.
- 07) Die Schätzungen der ungebundenen L1-Varianz (1,083) und der ungebundenen L2-Varianz (0,165) sind hochsignifikant, sodass in die Modellschätzung (unter rein statistischen Gesichtspunkten) noch weitere “erklärende” Prädiktoren aufgenommen werden könnten.
- 08) Die Schätzwerte für die freien (ungebundenen) L1- und L2-Varianzen sind identisch mit denjenigen von Modell 12. Da jedoch Modell 13 einen freien Parameter weniger zur Modellschätzung hat (13 vs. 12 freie Parameter) sollte Modell 13 gegenüber Modell 12 bevorzugt werden (wenn auch der Unterschied sehr gering ausfällt). Das wird auch von den AIC- und BIC-Werten der Modelle 12 und 13 bestätigt: die AIC-Werte sind fast identisch und nur der BIC-Wert wird von Modell 12 zu Modell 13 deutlich kleiner.
- 09) Viele Modellannahmen, die für eine ML-Schätzung des Modells 13 gelten sollten (z.B. Abwesenheit von Multikollinearität u.a.) wurden hier aus Zeitgründen nicht untersucht, sie müssten allerdings üblicherweise bei einer ML-Modellschätzung getestet werden (vgl. dazu Kap. A2).
- 10) Zur Inspektion der Signifikanz aller geschätzten Modellgrößen vgl. die folgende Abbildung. Die Markierung der Schätzwerte mit \*\*\* bedeutet, dass diese als hochsignifikant ( $p=0,000$ ) anzusehen sind. Eine schwache Signifikanz ist mit \* ausgezeichnet (z.B. 0,061\*). Bei einer Irrtumswahrscheinlichkeit  $> 0,05$  und damit bei einer nicht vorhandenen Signifikanz wird der entsprechende Schätzwert mit (ns) gekennzeichnet.

Abb. 13.3: Graphische Skizze des Gleichungssystems von Modell 13 mit Schätzergebnaten



Tab. 13.1: Ergebnisse der MLA-Schätzung von Modell 13: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	3,765	0,000
L2-Intercept, ( $\gamma_{10}$ ) Slope(PA2)	0,714	0,000
L2-Slope, ( $\gamma_{01}$ ) LEZU ON GDP2	0,458	0,000
L2-Slope2, ( $\gamma_{20}$ ) Mean, Slope2(SEX2)	0,139	0,000
L2-Slope1, ( $\gamma_{11}$ ) Slope1 ON GDP2	0,097	0,213

Tab. 13.2: Ergebnisse der MLA-Schätzung von Modell 13: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, Var( $e_{ij}$ )	1,083	0,000
L2-Residualvarianz, LEZU, Var( $u_{0j}$ )	0,165	0,000
L2-Residualvarianz, Slope1(PA2) Var( $u_{1j}$ )	0,033	0,007
L2-Varianz, Slope2(SEX2) Var( $u_{2j}$ )	0,012	0,176

Tab. 13.3: Ergebnisse der MLA-Schätzung von Modell 13: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
$\text{Cov}(u_{0j}, u_{1j})$	-0,008	0,518
$\text{Cov}(u_{0j}, u_{2j})$	-0,029	0,070
$\text{Cov}(u_{1j}, u_{2j})$	0,006	0,476

Tab. 13.4: Mplus-Inputfile für Modell 13

```

TITLE:
  Modell "MLA 13"
  mit L1-Prädiktor: "PA2" (dichotom, Mediansplit)
  mit L1-Prädiktor: "SEX2" (dichotom)
  mit L2-Prädiktor: "gdp2" (dichotom) (zur Schätzung von Intercepts + Slopes)
  mit random intercepts & mit random slopes

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_ME
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu ! lezu ist die abh. Y-Variable
  pa2 !
  sex2 !
  gdp2; !

WITHIN ARE pa2 sex2; ! pa2 + sex2 werden als L1-Prädiktoren definiert

BETWEEN ARE gdp2; ! definiert gdp2 als L2-Prädiktor (dichotomisiert)

CLUSTER=nation;

DEFINE:
  ! statt 1/2 nunmehr 0/1-Dichotomisierung
  IF(sex EQ 1) THEN sex2=0; ! 0=männl, 1=weibl
  IF(sex EQ 2) THEN sex2=1;

  ! hier kommt Mediansplit:
  IF(pa LE 5.67 )THEN pa2=0;
  IF(pa GT 5.67) THEN pa2=1;

  ! CENTER pa(GRANDMEAN); ! zentriert pa am Gesamtmittelwert
  ! keine Zentrierung bei Dummies

  ! hier kommt der Mediansplit:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74) THEN gdp2=1;

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
                           ! wenn random slopes geschätzt werden sollen

```

FORTSETZUNG -&gt;

Fortsetzung von Tabelle 13.4:

MODEL: %WITHIN%

Beta1j | lezu ON pa2 ; ! Spezifikation der L1-Regression mit Random Slope,  
! Beta1j ist Name für Random Slope

Beta2j | lezu ON sex2 ; ! L1-Regression

%BETWEEN%

lezu Beta1j ON gdp2; ! Schätzung der Intercepts und Slopes  
! durch L2-Prädiktor

lezu WITH Beta1j; ! Residual-Kovarianz zwischen  
! Random Intercepts und random slope1s

lezu WITH Beta2j;

Beta2j WITH Beta1j;

OUTPUT:

sampstat;

Tab. 13.5: Mplus-Ausgabe von Modellschätzung 13 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				12
Loglikelihood				
H0 Value				-10189.624
H0 Scaling Correction Factor for MLR				1.1519
Information Criteria				
Akaike (AIC)				20403.247
Bayesian (BIC)				20485.369
Sample-Size Adjusted BIC				20447.236
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Residual Variances				
LEZU	1.083	0.034	31.756	0.000
Between Level				
BETA1J ON				
GDP2	0.097	0.078	1.244	0.213
LEZU ON				
GDP2	0.458	0.118	3.882	0.000
LEZU WITH				
BETA1J	-0.008	0.013	-0.647	0.518
BETA2J	-0.029	0.016	-1.813	0.070
BETA2J WITH				
BETA1J	0.006	0.008	0.713	0.476
Means				
BETA2J	0.139	0.034	4.132	0.000
Intercepts				
LEZU	3.765	0.090	41.683	0.000
BETA1J	0.714	0.066	10.800	0.000
Variances				
BETA2J	0.012	0.009	1.354	0.176
Residual Variances				
LEZU	0.165	0.038	4.367	0.000
BETA1J	0.033	0.012	2.678	0.007



Modell 14:

Zwei-Ebenen-Modell,

Random-Intercepts-/Fixed+Random-Slopes-Modell (mit cross-level-Effekt),

mit L1-Prädiktor "PA2" (dichotomisiert, Mediansplit),

mit L1-Prädiktor "SEX2" (dichotomisiert),

mit L2-Prädiktor "GDP2"(dichotomisiert, Mediansplit)

zur Schätzung von Intercepts und Slopes

("intercepts-and-slopes-as-outcome model")

L1-Gleichung:  $Y_{ij} = \beta_{0j} + \beta_{1j}X1_{ij} + \beta_{2j}X2_{ij} + e_{ij}$  (14.1)

L2-Gleichungen:  $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$  (14.2a)

$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$  (14.2b)

$\beta_{2j} = \gamma_{20}$  (14.2.c)

"fixed part":  $\gamma_{00}, \gamma_{10}, \gamma_{20}, \gamma_{01}, \gamma_{11}$

"random part":  $\text{Var}(e_{ij}), \text{Var}(u_{0j}), \text{Var}(u_{1j})$

"covariance":  $\text{Cov}(u_{0j}, u_{1j})$

(Wie Modell 13, aber mit fixed SEX2-Effekten.)

Abb. 14.1: Graphische Skizze des Gleichungssystems von Modell 14

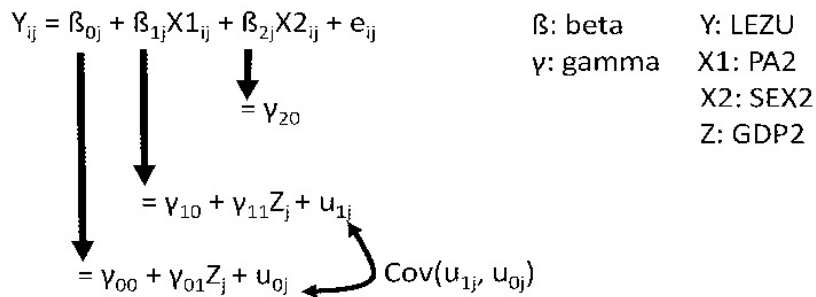
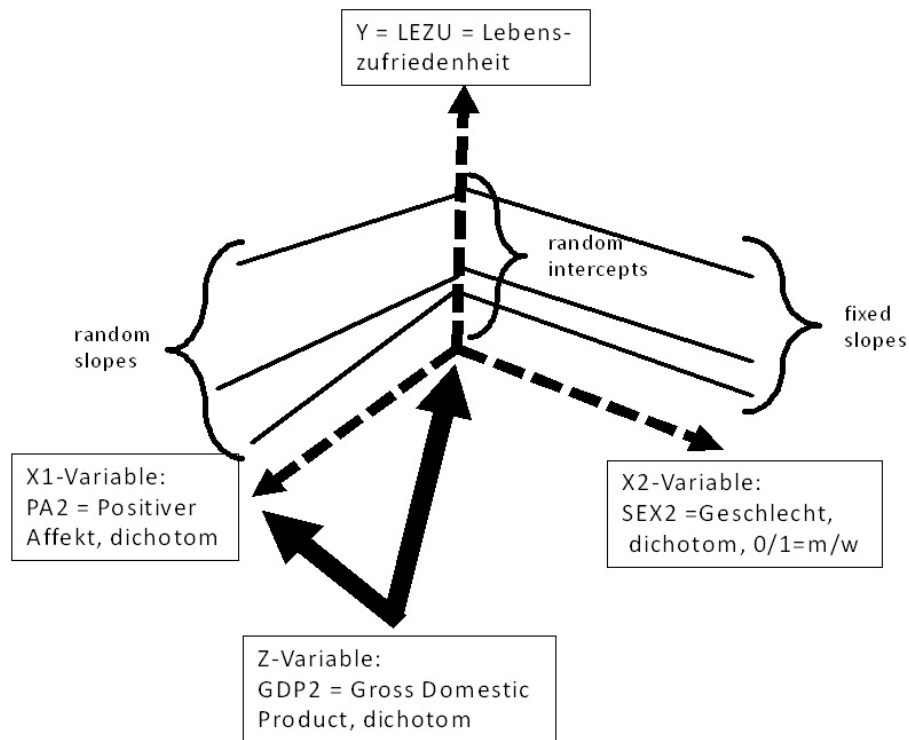


Abb. 14.2: Stilisierte Darstellung der Modell-Logik von Modell 14



#### Zur Logik von Modell 14:

Das Modell 14 unterscheidet sich nur dadurch von Modell 13, dass nunmehr der Effekt von SEX2 als “fixed effect” spezifiziert wird (in Modell 13 war dieser Effekt noch als “random effect” definiert worden). Es wird nunmehr also angenommen, dass der Effekt von SEX2 auf LEZU nicht zwischen den einzelnen Ländern variiert, sondern konstant ist. Es werden deshalb die ersten drei Gleichungen von Modell 13 auch für die Spezifikation von Modell 14 übernommen (14.1, 14.2a, 14.2b). Allein die Gleichung 13.2c wird durch die neue Gleichung 14.2c (s.o.) ersetzt.

Die Abbildung 14.2 verdeutlicht die Logik von Modell 14. Das Modell 14 ist weitgehend identisch mit Modell 13 (vgl. “Die Abbildung 13.2 ...”). Die beiden Modelle unterscheiden sich nur dadurch voneinander, dass das Modell 14 fixed SEX2-Effekte aufweist, die also nicht länderspezifisch variieren, sondern für alle Länder den gleichen Slope aufweisen (die durchgezogenen Linien zwischen Y- und X2-Koordinate haben alle die gleiche Steigung).

#### Zur Interpretation des Gleichungssystems von Modell 14:

Aufgrund der weitgehenden Strukturgleichheit der Modelle 12/13/14 wird im Folgenden allein die Gleichung 14.2c erläutert. Zum Verständnis der Gleichungen 14.1 bis 14.2b können die diesbezüglichen Ausführungen zu Modell 12 benutzt werden. Die einzelnen Parameterschätzwerte, die sehr ähnlich den Schätzwerten von Modell 12 und Modell 13 sind, können der folgenden Abbildung 14.3 sowie den Tabellen 14.1 bis 14.3 entnommen werden.

$$\beta_{2j} = \gamma_{20} \quad (14.2c)$$

Ist die Gleichung zur Schätzung eines fixen SEX2-Slopes,

$\gamma_{20}$  konstanter/fixer, länderübergreifender SEX2-Slope

Mplus: Within Level / LEZU ON SEX2 = 0,141 (p = 0,000)

#### Zur Interpretation des Schätzung von Modell 14:

(vgl. dazu auch die folgende Abbildung 14.2 sowie die Tabellen 14.1 bis 14.3)

- 01) Das Intercept der Gleichung zur Schätzung von LEZU bzw. der Durchschnittswert von LEZU beträgt dann, wenn alle Prädiktoren einen Wert von 0 aufweisen: 3,76 (LEZU-Skala: 1 - 7). Dies ist dann der Fall, wenn der dichotomisierte L1-Prädiktor PA2 (0/1) einen unterdurchschnittlichen Wert aufweist, wenn der dichotomisierte L1-Prädiktor SEX2 (0/1) die Ausprägung "männlich" annimmt und wenn der dichotomisierte L2-Prädiktor GDP2 (0/1) einen unterdurchschnittlichen Wert betrifft.
- 02) Das geschätzte mittlere Niveau von LEZU (=3,76, s.o.) erhöht sich signifikant um 0,45 Punkte, wenn die Befragten aus Ländern mit überdurchschnittlichen GDPs kommen.
- 03) Wenn PA2 von 0 auf 1 ansteigt (von unterdurchschnittlich auf überdurchschnittlich), vergrößert sich LEZU signifikant um 0,71 Punkte. Diese Punktzahl erhöht sich noch einmal um 0,09 Punkte, wenn GDP2 von 0 auf 1 ansteigt (von unterdurchschnittlich auf überdurchschnittlich). Dies wäre ein cross-level-Effekt. Allerdings ist der Zuschlag von 0,09 Punkten nicht signifikant und deshalb gibt es bezüglich des PA2-Prädiktors keinen von GDP2 ausgehenden cross-level-Effekt.
- 04) Wenn SEX2 von 0 auf 1 ansteigt (von männlich auf weiblich), vergrößert sich LEZU signifikant um 0,14 Punkte. Frauen haben also einen signifikant höheren LEZU-Wert als Männer. Dieser Effekt ist fixed d.h. er gilt für alle analysierten Länder.
- 05) Es gibt in Modell 13 keinen signifikanten, von GDP2 (Level 2) ausgehenden und auf PA2 ausgerichteten cross-level-Effekt.
- 06) Der Effekt von PA2 auf LEZU (0,71) ist hochsignifikant und bei GDP2=0 um ca. fünfmal stärker als der hochsignifikante Effekt von SEX2 auf LEZU (0,14). Der hochsignifikante Effekt von GDP2 auf LEZU (0,45) ist um ca. ein Drittel schwächer als der Effekt von PA2 (0,71). Es gilt also: Effekt(PA2) > Effekt(GDP2) > Effekt(SEX2). Dieser Vergleich ist möglich, weil alle L1- und L2-Prädiktoren dichotomisiert wurden.
- 07) Die Schätzungen der ungebundenen L1-Varianz (1,086) und der ungebundenen L2-Varianz (0,139) sind hochsignifikant, sodass in die Modellschätzung (unter rein statistischen Gesichtspunkten) noch weitere "erklärende" Prädiktoren aufgenommen werden könnten.

- 08) Die Schätzwerte für die freien (ungebundenen) L1- und L2-Varianzen sind sehr ähnlich denjenigen von Modell 13. Da jedoch Modell 14 drei freie Parameter weniger zur Modellschätzung hat als Modell 13 (12 vs. 9 freie Parameter) sollte Modell 14 gegenüber Modell 13 bevorzugt werden. Das wird auch durch einen Vergleich der AIC-/BIC-Werte von Modell 13 und Modell 14 bestätigt: die AIC-Werte sind weitgehend identisch (M13, AIC= 20403 / M14, AIC=20404), während der BIC-Wert von Modell 14 deutlich niedriger ist als der BIC-Wert von Modell 13 (M13, BIC=20485 / M14, BIC=20466) (vgl. auch Kap. B5).
- 09) Viele Modellannahmen, die für eine ML-Schätzung des Modells 14 gelten sollten (z.B. Abwesenheit von Multikollinearität u.a.) wurden hier aus Zeitgründen nicht untersucht, müssen allerdings bei einer qualitativ hochstehenden ML-Modellschätzung zusätzlich getestet werden (vgl. dazu Kap. A2).
- 10) Zur Inspektion der Signifikanz aller geschätzten Modellgrößen vgl. die folgende Abbildung. Die Markierung der Schätzwerte mit \*\*\* bedeutet, dass diese als hochsignifikant ( $p=0,000$ ) anzusehen sind. Eine schwache Signifikanz ist mit \* ausgezeichnet (z.B. 0,061\*). Bei einer Irrtumswahrscheinlichkeit  $> 0,05$  und damit bei einer nicht vorhandenen Signifikanz wird der entsprechende Schätzwert mit (ns) gekennzeichnet.

Abb. 14.3: Graphische Skizze des Gleichungssystems von Modell 14 mit Schätzergebnissen

$$\text{LEZU}_{ij} = \beta_{0j} + \beta_{1j}(\text{PA2}_{ij}) + \beta_{2j}(\text{SEX2}_{ij}) + e_{ij}$$

$\beta$ : beta      Y: LEZU  
 $\gamma$ : gamma    X1: PA2  
                   X2: SEX2  
                   Z: GDP2

$$= \gamma_{20} \quad (0,14^{***})$$

$$= \gamma_{10} + \gamma_{11}(\text{GDP2}_j) + u_{1j} \quad (0,71^{***}) \quad (0,09^{\text{ns}})$$

$$= \gamma_{00} + \gamma_{01}(\text{GDP2}_j) + u_{0j} \quad (3,76^{***}) \quad (0,45^{***})$$

-0,006(ns)

Tab. 14.1: Ergebnisse der MLA-Schätzung von Modell 14: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L2-Intercept, ( $\gamma_{00}$ ) LEZU	3,761	0,000
L2-Intercept, ( $\gamma_{10}$ ) Slope1 (PA2)	0,717	0,000
L1-Slope, ( $\gamma_{20}$ ) LEZU ON SEX2	0,141	0,000
L2-Slope, ( $\gamma_{01}$ ) LEZU ON GDP2	0,458	0,000
L2-Slope, ( $\gamma_{11}$ ) Slope1 ON GDP2	0,096	0,220

Tab. 14.2: Ergebnisse der MLA-Schätzung von Modell 14: “random effects coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, LEZU, $\text{Var}(e_{ij})$	1,086	0,000
L2-Residualvarianz, LEZU, $\text{Var}(u_{0j})$	0,139	0,000
L2-Residualvarianz, Slope1(PA2) $\text{Var}(u_{1j})$	0,034	0,008

Tab. 14.3: Ergebnisse der MLA-Schätzung von Modell 14: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
$\text{Cov}(u_{0j}, u_{1j})$	-0,006	0,614

Tab. 14.4: Mplus-Inputfile für Modell 14

```

TITLE:
  Modell "MLA  14"
  mit L1-Prädiktor: "PA2" (dichotom, Mediansplit)
  mit L1-Prädiktor: "SEX2" (dichotom)
  mit L2-Prädiktor: "gdp2" (dichotom, Mediansplit
    (zur Schätzung von Intercepts + Slopes)
  mit random intercept & mit random slope (PA2) & mit fixed slope (SEX2)

DATA:
  File is "E:/DATA/mplus/kultur.dat" ;

VARIABLE:
  NAMES ARE  NATION  L1_ID  SEX  AUF_E  KLA_E  AUF_F
             KLA_F  LEZU  PA  NA  BAL  IC
             GDP  AUF_E_ME  KLA_E_ME  AUF_F_ME  KLA_F_ME  LEZU_MEA
             PA_MEAN  NA_MEAN  BAL_MEAN  PA_ZEN  GDP_ZEN ;

MISSING ARE . ;

USEVARIABLES ARE
  nation ! nation ist die Cluster-Variable
  lezu  ! lezu ist die abh. Y-Variable
  pa2  !
  sex2 !
  gdp2;!

WITHIN ARE pa2 sex2; ! pa2 + sex2 werden als L1-Prädiktoren definiert

BETWEEN ARE gdp2; ! definiert gdp2 als L2-Prädiktor (dichotomisiert)

CLUSTER=nation;

DEFINE:
  ! statt 1/2 nunmehr 0/1-Dichotomisierung
  IF(sex EQ 1) THEN sex2=0; ! 0=männl, 1=weibl
  IF(sex EQ 2) THEN sex2=1;

  ! hier kommt Mediansplit:
  IF(pa LE 5.67 )THEN pa2=0;
  IF(pa GT 5.67) THEN pa2=1;

  ! hier kommt der Mediansplit:
  IF(gdp LE .74 )THEN gdp2=0;
  IF(gdp GT .74) THEN gdp2=1;

ANALYSIS:
  TYPE = TWOLEVEL random; ! Option "random" wird gebraucht,
                          ! wenn random slopes geschätzt werden sollen

```

FORTSETZUNG ->

Fortsetzung von Tabelle 14.4:

MODEL: %WITHIN%

Beta1j | lezu ON pa2 ; ! Spezifikation der L1-Regression mit Random Slope,  
! Beta1j ist Name für Random Slope

lezu ON sex2 ; ! fixed effect

%BETWEEN%

lezu Beta1j ON gdp2; ! Schätzung des Intercepts und Slopes  
! durch L2-Prädiktor

lezu WITH Beta1j; ! Residual-Kovarianz zwischen  
! Random Intercept und random slope1

OUTPUT:

sampstat;

Tab. 14.5: Mplus-Ausgabe von Modellschätzung 14 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				9
Loglikelihood				
H0 Value				-10193.268
H0 Scaling Correction Factor for MLR				1.2542
for MLR				
Information Criteria				
Akaike (AIC)				20404.535
Bayesian (BIC)				20466.126
Sample-Size Adjusted BIC				20437.526
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
LEZU ON				
SEX2	0.141	0.033	4.210	0.000
Residual Variances				
LEZU	1.086	0.034	31.791	0.000
Between Level				
BETA1J ON				
GDP2	0.096	0.078	1.226	0.220
LEZU ON				
GDP2	0.458	0.124	3.685	0.000
LEZU WITH				
BETA1J	-0.006	0.013	-0.505	0.614
Intercepts				
LEZU	3.761	0.094	39.928	0.000
BETA1J	0.717	0.067	10.720	0.000
Residual Variances				
LEZU	0.139	0.030	4.707	0.000
BETA1J	0.034	0.013	2.651	0.008



## **E Mehrebenen-Regression für 3 Ebenen**

### **E1 Die Gliederung**

Im Folgenden werden neun Drei-Ebenen-Regressionsmodelle vorgestellt. Wir gehen davon aus, dass diese und die daraus abgeleiteten Modelle zu den am häufigsten eingesetzten Modellen bei sozialwissenschaftlichen Drei-Ebenen-Regressionsanalysen gehören. Diese Modelle werden hier ausführlich erläutert, sodass ihre Logik leicht verständlich wird und gleichzeitig ihre praktische Spezifikation, ihre Schätzung und ihre Interpretation im Kontext von Mplus-Input/Output-Files leicht nachvollziehbar werden.

Jedes Modell wird im Folgenden in der gleichen Art und Weise vorgestellt und erläutert, sodass die Modelle leicht miteinander verglichen werden können. Zudem erfolgt zu Vergleichszwecken noch vor Erläuterung der einzelnen Modellschätzungen eine tabellarische Klassifikation aller neun Drei-Ebenen-Modelle (Kapitel E3).

Zu jedem der neun Modell werden Informationen in der folgenden, hier vorgestellten Reihenfolge vermittelt:

1. Benennung des Modells mit seinen wichtigsten Kennzeichen (die dafür benutzten Textzeilen sind unterstrichen);
2. Auflistung der formal-statistischen Gleichungen, die das Modell kennzeichnen;
3. Graphische Skizze des Gleichungssystems (Abbildung x.1);
4. Stilisierte graphische Darstellung der wichtigsten Kennzeichen des Modells und seiner zentralen Modell-Logik (Abbildung x.2);
5. Erläuterung der wichtigsten Modell-Logik und der Abbildung x.2 (vgl. Pkt. 4);
6. Interpretation des modellspezifischen Gleichungssystems mit seinen einzelnen Modellgleichungen (vgl. Pkt. 2);
7. Hinweise zur Interpretation einzelner Ergebnisse der Modellschätzung;
8. Graphische Skizze des Gleichungssystems (vgl. Pkt. 2) mit den wichtigsten Schätzergebnissen (Abbildung x.3);
9. Tabellen mit den Ergebnissen der MLA-Schätzung (fixed und random coefficients);
10. Tabelle mit Mplus-Inputfile und mit integrierten Kommentaren;
11. Tabelle mit Mplus-Outputfile und mit integrierten Kommentaren.

Die für die einzelnen Komponenten der MLA benutzte Schreibweise (Notation) wird in Kapitel 5 näher vorgestellt. Allerdings dürften unsere Ausführungen auch ohne intime Kenntnis der formal-statistisch korrekten Notation verständlich sein.

Alle neun Modellschätzungen benutzen den gleichen Datensatz. Diese Datensatz und die darin enthaltenen, analyserelevanten Variablen werden im folgenden Gliederungspunkt (E2) erläutert.

## E2 Der Datensatz

Zur Schätzung der neun Drei-Ebenen-Modelle 20 bis 28 wird der Datensatz:

ch4three.DAT

aus Heck/Thomas 2015:100f benutzt. Dieser Datensatz ist kostenfrei von der Homepage:

[www.routledge.com/9781848725522](http://www.routledge.com/9781848725522) (dort unter: Support Material, Datasets)

herunter zu laden.

Die Daten betreffen N=2868 befragte US-amerikanische Schüler:innen (Ebene: L1) in 250 Schulklassen (Ebene: L2) in 63 Grundschulen bzw. "elementary schools" (Ebene: L2). Die durchschnittliche L2-Clustergröße beträgt 11,47 Schüler:innen. Die durchschnittliche L3-Clustergröße beträgt 45,52 Schüler:innen.

Inhaltlich betrachtet können die Daten benutzt werden, um die Abhängigkeit der Mathematik-Kompetenz von verschiedenen Indikatoren der L2-Ebene "Schulklasse bzw. Klassenlehrer" und der L3-Ebene "Schule" zu untersuchen.

Die hier benutzten Variablen des Datensatzes "ch4three.DAT" sind:

abh. Variable:	Y: MATH	Ergebnis von Mathematiktests für jeden befragten Schüler / jede befragte Schülerin (472 ... 775 Punkte; Median=578; Mean=600,52)
unabh. Variablen:	X1: FEMALE (Ebene L1)	1: weiblich / 0: männlich (Mean=0,50; Missing: n=4)
	X2: LOWSES (Ebene L1)	personaler SES-Wert der Herkunftsfamilie von jeder/jedem befragten Schüler:in (SES = sozioökonomischer Status) 1: niedrig/gering; 0: anderes (Mean=0,38)
	Z1: TEFFECT (Ebene L2)	Bewertung jedes/jeder Klassenlehrer(s):in durch seine/ihre Schüler:innen hinsichtl. seiner/ihrer Lehrkompetenz ("teaching effectiveness") auf Klassenebene (standardisiert) (-2,48 ... 2,87; Mean= -0,03; Median= -0,09)
	Z2: LOWSES_M (Ebene L2)	Anteil von Schüler:innen mit SES=1 (niedrig, s.o.) in jeder Klasse (SES=sozioökonomischer Status), (0 ... 1; Mean=0,38; Median=0,33)
	W1: SCHSES_M (Ebene L3)	Anteil von Schüler:innen mit SES=1 (niedrig) pro Schule (SES=sozioökonomischer Status), Mean aller klassenspezifischen LOWSES_M (s.o.) einer jeden Schule (0 ... 1; Mean=0,37; Median=0,36)
	W2: TEFF_M (Ebene L3)	Bewertung aller Klassenlehrer:innen auf Schulebene, Mean von TEFFECT (s.o.) (-2,3 ... 1,35; Mean= -0,04; Median=0,01)
Cluster-IDs	C1: PERSON C2: TEACHID C3: SCHCODE	L1-Analyseebene: Schüler:in L2-Analyseebene: Klassenlehrer:in bzw. Schulklasse L3-Analyseebene: Schule (elementary school)

**E3 Tabellarische Klassifikation von 9 Drei-Ebenen-Modellen**

<b>Drei-Ebenen-Modelle:</b>		
	<b>Modelltyp:</b>	<b>Forschungsfrage(n):</b>
Modell 20	Random-Intercepts-Only-Modell -> ohne L1/2/3-Prädiktoren	Wie groß ist der Anteil der L2-Varianz bzw. der Anteil der L3-Varianz an der Gesamtvarianz von MATH? Sind die jeweiligen Anteile größer als 10% (Daumenregel), sodass eine MLA sinnvoll ist?
Modell 21	Random-Intercepts-/Fixed-Slopes-Modell -> mit je zwei L1/2/3-Prädiktoren -> intercepts-as-outcome model	Die Effekte der sechs Prädiktoren auf die variierenden MATH-Intercepts (=random intercepts) sollen kontextinvariant (=fixed slopes) geschätzt und (soweit es die jeweiligen Skalierungen erlauben) miteinander verglichen werden.
Modell 22	Random-Intercepts-/Fixed+Random-Slopes-Modell -> mit je zwei L1/2/3-Prädiktoren -> intercepts-as-outcome model	... wie Modell 21, aber mit random TEFFECT-Slope. Gefragt wird, ob der Effekt von TEFFECT (Bewertung des Klassenlehrers bzw. der Klassenlehrerin) auf MATH zwischen den Schulen signifikant variiert
Modell 23	Random-Intercepts-/Fixed+Random-Slopes-Modell -> mit je zwei L1/2/3-Prädiktoren -> intercepts-and-slopes-as-outcome model -> cross-level-effects	... wie Modell 22, aber mit den beiden L3-Prädiktoren "W1/W2" zur Schätzung des TEFFECT-Slopes (als abh. L2-Var.). Gefragt wird, ob die Varianz des L2-TEFFECT-Slope mit den Effekten der zwei L3-Prädiktoren zu "erklären" ist
Modell 24	Random-Intercepts-/Fixed+Random-Slopes-Modell -> mit je zwei L1/2/3-Prädiktoren -> intercepts-and-slopes-as-outcome model -> cross-level-effects -> dichotome unabh. Prädiktoren	... wie Modell 23, aber alle unabh. L1/2/3-Prädiktoren sind nun dichotomisiert (teilw. durch Mediansplit 0/1), sodass die Effektstärken miteinander verglichen werden können
Modell 25	Random-Intercepts-/Fixed+Random-Slopes-Modell -> mit je zwei L1/2/3-Prädiktoren -> intercepts-as-outcome model -> dichotome unabh. Prädiktoren	... wie Modell 21, aber alle unabh. L1/2/3-Prädiktoren sind nun dichotomisiert (teilw. durch Mediansplit 0/1), sodass die Effektstärken miteinander verglichen werden können. Gefragt wird, ob der nunmehr nicht mehr fixierte L1-Slope " $\beta_{2jk}$ " auf den L2 u. L3-Ebenen variiert, sodass er im Folgemodell 26 als abh. Variable untersucht werden kann
Modell 26	Random-Intercepts-/Fixed+Random-Slopes Modell -> mit je zwei L1/2/3-Prädiktoren -> dichotome unabh. Prädiktoren -> intercepts-and-slopes-as-outcomes model -> cross-level-effect (Prädiktor Z2)	... wie Modell 25, allerdings ist der LOWSES-Slope jetzt eine abh. Var., die vom cross-level-effect von LOWSES_M beeinflusst wird. Damit soll getestet werden, ob der L1-Slope von LOWSES vom L2-Prädiktor LOWSES_M modifiziert werden kann

<p>Modell 27</p>	<p>Random-Intercepts-/Fixed Slopes Modell          -&gt; mit je zwei L1/2/3-Prädiktoren          -&gt; dichotome unabh. Prädiktoren          -&gt; intercepts-as-outcomes model          -&gt; mit L1-Interaktionsvariable (X1*X2)</p>	<p>... wie Modell 21, allerdings sind alle Prädiktoren nunmehr dichotom. Gefragt wird, ob Y von der zusätzlichen Interaktionsvariablen X1*X2 beeinflusst wird</p>
<p>Modell 28</p>	<p>Random-Intercepts- / Fixed + Random Slopes Modell          -&gt; mit je zwei L1/2/3-Prädiktoren          -&gt; dichotome unabh. Prädiktoren          -&gt; intercepts-as-outcomes model          -&gt; mit L1-Interaktionsvariable (X1*X2)          -&gt; freie Variation von X1*X2 auf L2/3</p>	<p>... wie Modell 27, allerdings kann X1*X2 jetzt frei auf L2/3 variieren. Gefragt wird, ob X1*X2 auf L2/3 bedeutsam bzw. signifikant variiert</p>

**E4 Systematische Erläuterung von 9 Drei-Ebenen-Modellen**

(Modell 20 - Modell 28)

Modell 20: Nullmodell, ohne Prädiktoren. . . . .	154
(random-intercepts-only model)	
Modell 21: fixe Prädiktoren-Effekte auf Intercepts . . . . .	161
(intercepts-as-outcome model)	
Modell 22: wie M21, aber mit Random-TEFFECT-Slopes . . . . .	170
(intercepts-as-outcome model)	
Modell 23: wie M22, aber mit Prädiktoren zur Schätzung des Random-Slopes. . . . .	178
(intercepts-and-slopes-as-outcome model, cross-level-effects)	
Modell 24: wie M23, aber alle Prädiktoren nunmehr dichotomisiert . . . . .	185
(intercepts-and-slopes-as-outcome model, cross-level-effects)	
Modell 25: wie M21, aber alle Präd. dichotomisiert und mit Random-X2-Effekt . . .	193
(intercepts-and-slopes-as-outcome model, cross-level-effects)	
Modell 26: wie M25, aber mit abh. Random-X2-Effekt. . . . .	200
(intercepts-and-slopes-as-outcome model, cross-level-effects)	
Modell 27: wie M21, aber mit dichotomen Präd. und fixem $X1 * X2$ -Effekt . . . . .	207
(intercepts-as-outcome model)	
Modell 28: wie M27, aber mit random $X1 * X2$ -Effekt . . . . .	215
(intercepts-as-outcome model)	

Modell 20:

Drei-Ebenen-Modell,

Random-Intercepts-Only-Modell,

ohne L1- /L2-/ L3-Prädiktoren

$$\text{L1-Gleichung: } Y_{ijk} = \beta_{0jk} + e_{ijk} \quad (20.1)$$

$$\text{L2-Gleichung: } \beta_{0jk} = \gamma_{00k} + u_{0jk} \quad (20.2a)$$

$$\text{L3-Gleichung: } \gamma_{00k} = \delta_{000} + v_{00k} \quad (20.3a)$$

“fixed part”:  $\delta_{000}$

“random part”:  $\text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(v_{00k})$

(Das Modell 20 entspricht dem Modell 1 und dessen Formalisierung in Heck/Thomas 2015: 101-103.)

Abb. 20.1: Graphische Skizze des Gleichungssystems von Modell 20

$$\begin{array}{l}
 Y_{ijk} = \beta_{0jk} + e_{ijk} \\
 \downarrow \\
 = \gamma_{00k} + u_{0jk} \\
 \downarrow \\
 = \delta_{000} + v_{00k}
 \end{array}
 \quad
 \begin{array}{l}
 \beta: \text{beta} \\
 \gamma: \text{gamma} \\
 \delta: \text{delta} \\
 Y: \text{math}
 \end{array}$$

Zur Logik von Modell 20:

Das Random-Intercepts-Only-Modell (auch “Null-Modell” oder “Baseline-Modell” genannt) hat keine Prädiktoren. Es werden allein durchschnittliche Y-Werte geschätzt. Das Modell dient zur Aufteilung der Gesamtvarianz von kontextuell eingebetteten Y-Messungen in eine L1-Varianz (“within” Varianz auf der Individualebene) und in eine L2-Varianz (“between” Varianz auf den Kollektivebenen L2 und L3) (vgl. dazu Kap. A1). Die Aufteilung der Gesamtvarianz in einzelne Teilvarianzen ist konstitutiv für die MLA. Mit den Teilvarianzen lassen sich die (problematischen) Intra-Klassen-Korrelationskoeffizienten (ICC-Koeffizienten) auf der L2- und der L3-Ebene berechnen (vgl. Kap. A1). Die ICCs geben an, wie groß der prozentuale Anteil der L2- und der L3-Varianz an der Gesamtvarianz einer abh. Variablen ist. Nach einer häufig benutzten (problematischen) Daumenregel sollte der Anteil der L2- bzw. L3-Varianz mindestens 10% an der Gesamtvarianz betragen, damit eine MLA deutliche Vorteile gegenüber einer klassischen Ein-Ebenen-Analyse aufweist.

Die ICCs berechnen sich nach den folgenden Formeln 20.4a und 20.4b. Sie werden aber auch direkt von Mplus ausgegeben (vgl. im Folgenden: “Zur Interpretation der Schätzung ...” Pkt. 2).

$$ICC_{L2} = \text{Varianz}_{L2} / (\text{Varianz}_{L1} + \text{Varianz}_{L2} + \text{Varianz}_{L3}) \quad (20.4a)$$

$$ICC_{L3} = \text{Varianz}_{L3} / (\text{Varianz}_{L1} + \text{Varianz}_{L2} + \text{Varianz}_{L3}) \quad (20.4b)$$

Mit der Schätzung von Modell 20 soll ermittelt werden:

- a) wie hoch die L2- und L3-ICC-Koeffizienten sind,
- b) ob die Anteile der beiden Between-Varianzen an der Gesamtvarianz groß genug sind (>10%), um die vorliegenden Daten mittels einer Drei-Ebenen-Analyse und mit L2- sowie L3-Prädiktoren durchzuführen.

Zur Interpretation des Gleichungssystems von Modell 20:

$$Y_{ijk} = \beta_{0jk} + e_{ijk} \quad (20.1)$$

Als abh. Variable "Y<sub>ijk</sub>" auf der Mikroebene L1 dient MATH. Der MATH-Wert gilt jeweils für jeden/jede i-ten Schüler:in in der j-ten Klasse an der k-ten Schule. Für jeden/jede Schüler:in wird ein MATH-Wert geschätzt, der identisch mit dem MATH-Mittelwert aller Schüler:innen einer jeweiligen Schulklasse/Schule ist. Dieser mit  $\beta_{0jk}$  geschätzte Y-Wert entspricht fast niemals dem individuellen, tatsächlichen Y-Wert der Schüler:innen. Der reale Wert wird oberhalb oder unterhalb des geschätzten Mittelwerts liegen. Die Differenz zwischen dem geschätzten und dem empirischen Y-Wert wird als "e" (L1-Residuum oder Residualvariable) bezeichnet. Als "Var(e<sub>ijk</sub>)" wird die Varianz der Residualvariablen e<sub>ijk</sub> auf der L1-Ebene bezeichnet. Sie wird auch "Within-Varianz" genannt (Varianz innerhalb der einzelnen Schulklassen).

Mplus: Within Level / Variances MATH: 1454,394 (p=0,000)

$$\beta_{0jk} = \gamma_{00k} + u_{0jk} \quad (20.2a)$$

Über alle einzelnen, klassenspezifischen Y-Mittelwerte lässt sich auch ein Gesamtmittelwert (auch: "Grand Mean" genannt) berechnen ( $\gamma_{00k}$ ). Die Differenzen zwischen diesem Grand Mean und den einzelnen Klassen-Mittelwerten ergeben die Werte der Residualvariablen "u<sub>0jk</sub>" auf der L2-Ebene "Schulklasse". Die Werte dieser Residualvariablen entsprechen den Abweichungen eines jeden klassenspezifischen Mittelwerts " $\beta_{0jk}$ " vom Gesamtmittelwert " $\gamma_{00k}$ ". Die Varianz der Residualvariablen "Var(u<sub>0jk</sub>)" liegt auf der L2-Ebene. Sie wird auch als Between-Varianz bezeichnet (Varianz zwischen den einzelnen Schulklassen).

Mplus: Between TEACHID Level / Variances MATH = 149,551 (p=0,004)

$$\gamma_{00k} = \delta_{000} + v_{00k} \quad (20.3a)$$

Ein Gesamtmittelwert ("Grand Mean"  $\delta_{000}$ ) kann auch über alle schulspezifischen Y-Mittelwerte ermittelt werden. Im Gleichungssystem 20.1 bis 20.3a ist der Grand Mean " $\delta_{000}$ " der einzige feste Parameter (fixed coefficient), der einen bestimmten Wert aufweist und nicht variiert (Eselsbrücke: es gibt für ihn keine Subskripte i, j, oder k).

Mplus: Between SCHCODE Level / Means MATH = 598.616 (p=0,000)

Die Residualvariable "v<sub>00k</sub>" liegt auf der L3-Ebene "Schule". Der Variablenwert entspricht der Abweichung eines jeden schulspezifischen Mittelwerts " $\gamma_{00k}$ " vom Gesamtmittelwert aller Schulmittelwerte " $\delta_{000}$ ". Die Varianz "Var(v<sub>00k</sub>)" der Residualvariablen v<sub>00k</sub> wird auch als Between-Varianz bezeichnet (Varianz zwischen den Schulen).

Mplus: Between SCHCODE Level / Variances MATH = 226,518 (p=0,001)



Zur Interpretation der Schätzung von Modell 20:

(vgl. dazu auch die folgende Abbildung 20.2 sowie die folgenden Tabellen 20.1 und 20.2)

- 1) Die Residualvariablen  $e_{ijk}$  (L1-Ebene) und  $u_{0jk}$  (L2-Ebene) sowie  $v_{ijk}$  (L3-Ebene) können als zufällige Größen (random coefficients) bezeichnet werden, da sie mit Auswahl eines jeweiligen Befragten “i” und/oder einer jeweiligen Schulklasse “j” und/oder einer jeweiligen Schule “k” variieren (Eselsbrücke: es gibt für die Variablen auch Subskripte: i, j und/oder k). Im Unterschied dazu ist  $\gamma_{000}$  in den Gleichungen des Modells der einzige feste Parameter (fixed coefficient), der einen bestimmten Wert aufweist und der auch nicht variieren kann (Eselsbrücke: es gibt für ihn kein Subskript i, j und/oder k).
  
- 2) Mit den Varianzen aus dem Random-Intercepts-Only-Modell und den Gleichungen 20.4a sowie 20.4b können die (problematischen) L2- und L3-Intra-Klassen-Korrelationen (ICC) berechnet werden (vgl. Kap. A1). Allerdings werden die ICCs auch direkt vom Mplus-Output ausgegeben:  
Mplus: SUMMARY OF DATA / Estimated Intraclass Correlations for the Y Variables for TEACHID level / Variable MATH: 0,082  
Mplus: SUMMARY OF DATA / Estimated Intraclass Correlations for the Y Variables for SCHCODE level / Variable MATH: 0,124
  
- 3) Entsprechend der o.g. Ausgabe beträgt der Anteil der L2-Varianz an der Gesamtvarianz von Y ca. 8,2% und der Anteil der L3-Varianz an der Gesamtvarianz ca. 12,4%. Somit ist die Clusteraufteilung der Daten nach Schulklassen bzw. nach Schulen für 8,2% bzw. für 18,5% der Gesamtvarianz im Modell verantwortlich. Beide ICCs liegen von daher dicht an dem (problematischen) Grenzwert ( $ICC > 10\%$ ), ab dem nach einer Daumenregel eine MLA zur Auswertung von Beziehungen zwischen Mehr-Ebenen-Variablen vorgenommen werden sollte (vgl. Kap. A1). Im Kontext von MLAs könnte damit auch auf allen Analyseebenen nach solchen Prädiktoren gesucht werden, welche die freien Varianzanteile binden können.

Abb. 20.2: Graphische Skizze des Gleichungssystems von Modell 20 mit Schätzergebnaten

$$\begin{aligned}
 Y_{ijk} &= \beta_{0jk} + e_{ijk} \\
 &\downarrow \\
 &= \gamma_{00k} + u_{0jk} \\
 &\downarrow \\
 &= \delta_{000} + v_{00k} \\
 &\quad (598,61^{***})
 \end{aligned}$$

$\beta$ : beta  
 $\gamma$ : gamma  
 $\delta$ : delta  
 $Y$ : math

Tab. 20.1: Ergebnisse der MLA-Schätzung von Modell 20: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L3-Intercept ( $\delta_{000}$ ) Mean (MATH)	598,616	0,000

Tab. 20.2: Ergebnisse der MLA-Schätzung von Modell 20 : “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Varianz: MATH, $\text{Var}(e_{ijk})$	1454,394	0,000
L2-Varianz: MATH, $\text{Var}(u_{0jk})$	149,551	0,004
L3-Varianz: MATH, $\text{Var}(v_{00k})$	226,518	0,00

Tab. 20.3: Mplus-Inputfile für Modell 20:

```
TITLE:
  Model 20
  Three-level (null) model
  keine Prädiktoren (nur intercepts)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowses teffect
    lowses_m schses_m teff_m;
  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math;   ! ist die abh. Y-Variable
  CLUSTER = schcode teachid;
  MISSING ARE . ;

  WITHIN = ;   ! es gibt keinen kein L1-Prädiktor
  BETWEEN = ; ! es gibt keine L2 oder L3-Prädiktor

ANALYSIS:
  TYPE = THREELEVEL;

MODEL:
  %WITHIN%
    math;   ! Within-Varianz von MATH wird angefordert (innerhalb der Cluster)

  %BETWEEN teachid%
    math;   ! Between-Varianz von MATH wird angefordert
            ! (zwischen den teachid-Clustern)

  %BETWEEN schcode%
    math;   ! Between-Varianz von MATH wird angefordert
            ! (zwischen den schcode Clustern)

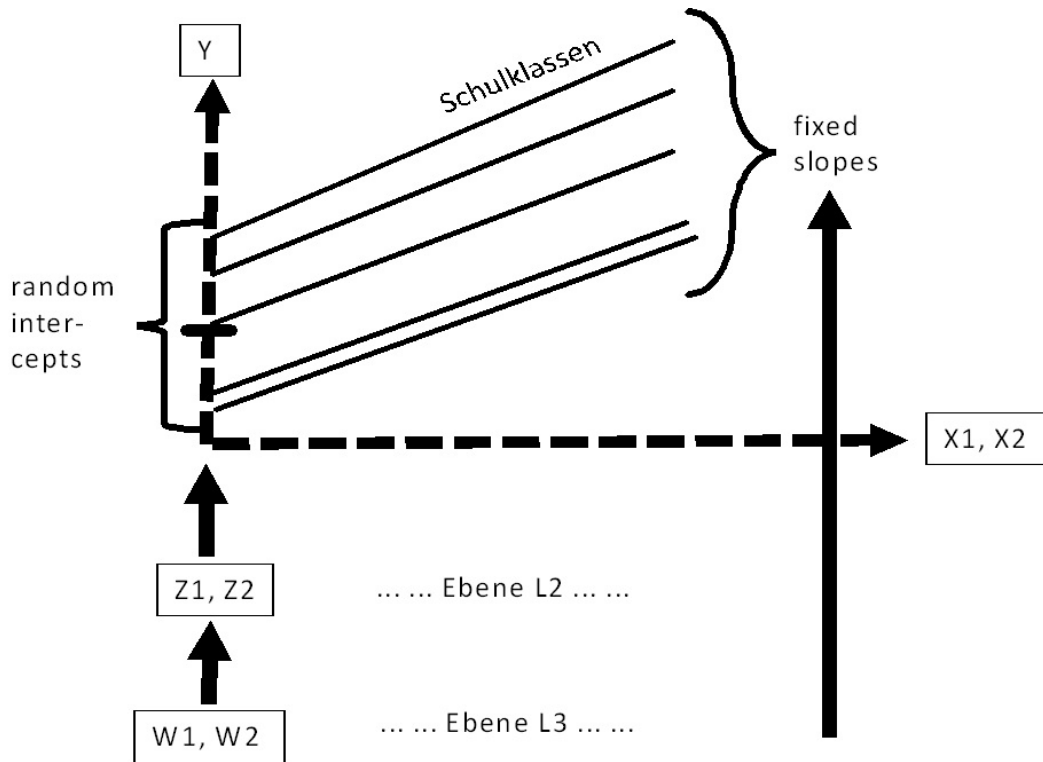
OUTPUT:
  SAMPSTAT;
  ! ACHTUNG: es wird eine Warnung ausgegeben,
  ! weil MATH nicht korreliert mit TEACHID oder SCHCODE,
  ! diese Warnung ist aber für unsere Zwecke irrelevant
```

Tab. 20.4: Mplus-Ausgabe von Modellschätzung 20 (ausgewählte Bereiche)

Average cluster size for TEACHID level		11.472		
Estimated Intraclass Correlations for the Y Variables for TEACHID level				
Intraclass				
Variable	Correlation			
MATH	0.082			
Average cluster size for SCHCODE level		45.524		
Estimated Intraclass Correlations for the Y Variables for SCHCODE level				
Intraclass				
Variable	Correlation			
MATH	0.124			
MODEL FIT INFORMATION				
Number of Free Parameters	4			
Loglikelihood				
H0 Value	-14636.913			
H0 Scaling Correction Factor for MLR	2.0762			
H1 Value	-14636.913			
H1 Scaling Correction Factor for MLR	2.0762			
Information Criteria				
Akaike (AIC)	29281.826			
Bayesian (BIC)	29305.672			
Sample-Size Adjusted BIC	29292.962			
(n* = (n + 2) / 24)				
MODEL RESULTS				
		Two-Tailed		
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
Variances				
MATH	1454.394	78.160	18.608	0.000
Between TEACHID Level				
Variances				
MATH	149.551	52.453	2.851	0.004
Between SCHCODE Level				
Means				
MATH	598.616	2.384	251.124	0.000
Variances				
MATH	226.518	65.121	3.478	0.001



Abb. 21.2: Stilisierte Darstellung der Modell-Logik von Modell 21



#### Zur Logik von Modell 21:

Das Modell 21 ist ein Random-Intercepts- /Fixed-Slopes-Modell, das über drei Analyseebenen reicht.

Das Modell enthält für jede der drei Analyseebenen je zwei Prädiktoren. Deren Effekte variieren nicht, sondern sind klassen- bzw. schulübergreifend gültig. So gelten der X1- und der X2-Effekt in gleicher Weise für alle Schüler:innen (L1), in jeder Schulklasse (L2) und in jeder Schule (L3). Und auch der Z1- sowie der Z2-Effekt gelten für alle Schulklassen in allen Schulen, wie auch der W1- und der W2-Effekt für alle Schulen gilt. Alle genannten Effekte werden im Modell auf ihrer jeweiligen Analyseebene konstant gesetzt und variieren nicht.

Während also die Slopes im Modell nicht variieren, variieren die personalen Y-Werte und die klassenspezifischen Y-Intercepts. Jeder/jede Schüler:in hat ihren/seinen speziellen Y-Wert und jede Schulklasse hat ihr spezielles Y-Intercept. So können diese Werte als abhängige Variablen zur Schätzung der Z- und W-Effekte dienen.

Mit dem Modell 21 können die Effektstärken der sechs Prädiktoren geschätzt und (soweit es die jeweiligen Skalierungen erlauben) die Effektstärken miteinander verglichen werden.

Die Abbildung 21.2 verdeutlicht die Logik von Modell 21. Das dargestellte Koordinatensystem (gestrichelte Linien) besteht aus der senkrechten Y-Achse (für die Variable MATH) und der waagrechten X1- bzw. X2-Achse (für die Variablen FEMALE bzw. LOWSES). Der Zusammenhang zwischen Y- und X-Variable wird für jede Schulklasse durch die Steigungslinie innerhalb des Koordinatensystems symbolisiert. Alle Steigungslinien haben die gleiche Steigungsrate (fixed slopes), aber für jede Schulklasse unterschiedlich große Intercepts (random intercepts). Ihre Ausprägungen werden von den beiden Z-Variablen und den beiden W-Variablen beeinflusst.

Zur Interpretation des Gleichungssystems von Modell 21:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + e_{ijk} \quad (21.1)$$

$$\beta_{1jk} = \gamma_{10k} \quad (21.2b)$$

$$\beta_{2jk} = \gamma_{20k} \quad (21.2c)$$

$$\gamma_{10k} = \delta_{100} \quad (21.3d)$$

$$\gamma_{20k} = \delta_{200} \quad (21.3e)$$

Mit den Daten der Schüler:innen wird für jede Schulklasse eine Lineargleichung zur Prognose der personalen MATH-Testergebnisse ( $Y=MATH$ ) geschätzt (Gl. 21.1). Als Prädiktoren dienen die dichotomen Variablen X1 (FEMALE=weibl. Geschlecht) und X2 (LOWSES = unterdurchschnittl. personaler SES-Wert). Die Effektstärken ergeben sich aus den geschätzten Slopes. Diese Slopes gelten sowohl über alle Schulklassen als auch über alle Schulen hinweg (fixed slopes). Diese Festlegung erfolgt für X1/FEMALE durch die Gleichungen Gl. 21.2b und Gl. 21.3d sowie für X2/LOWSES durch die Gleichungen Gl. 21.2c und Gl. 21.3e. Das Intercept “ $\beta_0$ ” berichtet den Y-Schätzwert für eine fiktiv angenommene Referenzgruppe, bei der alle Prädiktoren einen Wert von 0 annehmen. Das geschieht bei Schüler:innen mit X1/FEMALE=0 und LOWSES=0 sowie mit Klassenprädiktoren (Z1, Z2) und Schulprädiktoren (W1, W2) von 0 (wobei ein Wert von 0 bei diesen zentrierten Variablen identisch mit dem Durchschnittswert ist).

Die Abweichungen der geschätzten Y-Werte von den tatsächlichen Y-Werten ergeben die Residuen “e”. Aus den Residualwerten lässt sich die ungebundene (nicht “erklärte”) Within-Varianz berechnen.

Mplus: Within Level / MATH ON LOWSES: -13.674 (p=0,000)

Mplus: Within Level / MATH ON FEMALE: 4,073 (p=0,001)

Mplus: Within Level / Residual Variances MATH: 1419.607 (p=0,000)

Mplus: Between SCHCODE Level / Intercepts MATH: 601,865 (p=0,000)

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk} \quad (21.2a)$$

$$\gamma_{01k} = \delta_{010} \quad (21.3b)$$

$$\gamma_{02k} = \delta_{020} \quad (21.3c)$$

Die Intercepts aus der Linearschätzung nach Gl. 21.1 können als klassenspezifische  $Y=MATH$ -Werte (für die Referenzgruppe, s.o.) gelesen werden. Sie dienen in Gl. 21.2a als abhängige Größe für eine zweite Linearschätzung mit den klassenspezifischen Prädiktoren Z1/TEFFECT (aggregierte Bewertung des/der Klassenlehrer:in hinsichtlich seiner/ihrer Lehrkompetenz) und Z2/LOWSES\_M (Anteil der Schüler:innen mit unterdurchschnittlichem SES-Wert). Die Slopes markieren wiederum die Effektstärken der beiden Prädiktoren. Dabei sind die Slopes von Z1/TEFFECT und Z2/LOWSES\_M fixiert und gelten für alle Schulklassen wie auch für alle Schulen (nach Gl. 21.3b und Gl. 21.3c). Das Intercept berichtet den klassendurchschnittlichen  $\beta_{0jk}$ -Wert für die Referenzgruppe mit Werten von 0 bei Z1/TEFFECT, Z2/LOWSES\_M, W1/SCHSES\_M und W2/TEFF\_M (was aufgrund der Zentrierung der vier Variablen dem jeweiligen Durchschnittswert entspricht). Das Residuum “u” informiert über die Differenzen zwischen den nach Gl. 21.1 geschätzten klassenspezifischen Intercept-Werten und den schulspezifisch nach Gl. 21.2a geschätzten Intercepts, wobei die aus den Residuen berechnete Varianz die ungebundene (nicht “erklärte”) Between-Varianz auf der L2-Ebene ist.

Mplus: Between TEACID Level / MATH ON LOWSES\_M: -14.547 (p=0,054)

Mplus: Between TEACID Level / MATH ON TEFFECT: 7,218 (p=0,000)

Mplus: Between TEACID Level / Residual Variances MATH: 68.604 (p=0,006)

$$\gamma_{00k} = \delta_{000} + \delta_{001} W1_k + \delta_{002} W2_k + v_{00k} \quad (21.3a)$$

Die Gleichung 21.3a ist die dritte zu schätzende Lineargleichung im Modell 21. Geschätzt wird das zu erwartende Intercept “ $\gamma_{00k}$ “ aus Gl. 21.3a. Die Prädiktoren für die Schätzung sind die schulspezifischen L3-Variablen “SCHSES\_M” (Durchschnitt der klassenspezifischen Anteile von Schülern/Schülerinnen mit unterdurchschnittlichem SES-Wert pro Schule) und “TEFF\_M” (Durchschnitt der klassenspezifischen Bewertungen aller Lehrer:innen pro Schule). Die Interpretation der Gleichungskomponenten erfolgt in gleicher Weise wie oben bei Gl. 21.1 und 21.2a erläutert. Die Lineargleichung liefert auch den Schätzwert für das Intercept “ $\delta_{000}$ “. Die aus den Residuen berechnete Varianz ist die ungebundene (nicht “erklärte”) Between-Varianz auf der L3-Ebene.

Mplus: Between SCHCODE Level / MATH ON SCHSES\_M: -22.991 (p=0,021)

Mplus: Between SCHCODE Level / MATH ON TEFF\_M: 13,987 (p=0,000)

Mplus: Between SCHCODE Level / Intercepts MATH: 601,865 (p=0,000)

Mplus: Between SCHCODE Level / Residual Variances MATH: 23.171 (p=0,031)

#### Zur Interpretation der Schätzung von Modell 21:

(vgl. dazu auch die folgende Abbildung 21.3 sowie die folgenden Tabellen 21.1 und 21.2)

- 1) Lehrer:innen haben einen Einfluss auf die MATH-Fähigkeit von Schülern und Schülerinnen. Das gilt hinsichtlich der Bewertung der einzelnen Klassenlehrer:in (TEFFECT=7,21\*\*\*) als auch hinsichtlich der Bewertung aller Lehrer:innen einer Schule (TEFF\_M=13,98\*\*\*).
- 2) Ein unterdurchschnittlicher SES verringert auf allen Analyseebenen (Schüler:innen, Klassen, Schulen) die MATH-Fähigkeit der einzelnen Schüler:innen (LOWSES=-13,67\*\*\*; LOWSES\_M=-14,54<sup>ns</sup>; SCHSES\_M=-22,99\*), wenn auch der einzige hochsignifikante (und negative) Effekt auf der Schüler:innen-Ebene liegt. Unabhängig vom Signifikanzniveau betrachtet, wirkt jeder SES-Prädiktor für sich alleine und wird nicht von der Effektstärke der anderen beiden SES-Prädiktoren beeinflusst. Wenn also z.B. zwei Schüler:innen einen LOWSES=1 aufweisen, wirkt LOWSES\_M (Klasse) noch insofern, als jede Steigerung des LOWSES\_M-Anteils um eine Messeinheit die zu erwartende MATH-Leistung um 14,54 Punkte reduziert.
- 3) Schülerinnen (weibl.) haben bei Konstanz aller anderen Prädiktoren einen um ca. 4 Punkte höheren MATH-Leistungsstandard als Schüler (männl.).



- 4) Auf der Schüler:innen-Ebene ist der Effekt von LOWSES ca. dreimal so stark wie der Effekt von FEMALE (-13,67 vs. 4,07), allerdings haben beide Prädiktoren unterschiedliche Wirkungsrichtungen: Ein Wert von LOWSES=1 reduziert die MATH-Leistung, während ein Wert von FEMALE=1 die MATH-Leistung erhöht. Beide X-Effekte sind vergleichbar, weil die Prädiktoren gleich skaliert sind (mit 1/0).
- 5) Die Erweiterung von Modell 20 durch je zwei Prädiktoren auf jeder Analyseebene reduziert die freie Varianz des Modells insbesondere auf der L2- und der L3-Ebene ganz beträchtlich (L2: 149,55→68,60; L3: 226,51→23,17). Die diesbezüglichen Pseudo-R<sup>2</sup>-Werte betragen 0,54 bzw. 0,89 und besagen, dass die freie L2-Varianz durch Berücksichtigung von zwei Z-Variablen um 54% reduziert werden konnte, und dass die freie L3-Varianz durch Berücksichtigung von zwei W-Variablen um 89% reduziert werden konnte.<sup>29</sup>
- 6) Die Fit-Werte von Modell 20 und Modell 21 können nicht mit dem üblichen Chi-Quadrat-Differenzentest verglichen werden, da im vorliegenden Skript die Schätzung beider Modelle mit dem robusten ML-Schätzverfahren “MLR” zur Korrektur von Verstößen gegen die Verteilungsannahmen des ML-Verfahrens erfolgt (die MLR-Methode wird von Mplus als Voreinstellung gewählt).<sup>30</sup> Allerdings können zum Vergleich der Modellschätzungen (auch bei nicht-geschachtelten Modellen) die AIC- und die BIC-Werte ordinal (!) interpretiert werden: kleinere Werte indizieren dann bessere Schätzungen. Demnach ist die Schätzung von Modell 21 deutlich erfolgreicher als die Schätzung von Modell 20: AIC(M20)=29281,826 > AIC(M21)=29096,901; BIC(M20)=29305,672 > BIC(M21)=29156,515 (vgl. dazu die AIC- und BIC-Werte in Tab. 20.4 und Tab. 21.4) (vgl. auch Kap. B5).
- 7) Zur Interpretation einer jeden MLA-Modellschätzung mittels ML-Schätzverfahren gehört auch die Überprüfung der wichtigsten Modellannahmen. Zu diesen Annahmen gehören u.a.: das Vorhandensein einer ausreichenden Fallzahl, die Linearität der Variablenbeziehungen, die Abwesenheit von Multikollinearität, die multivariate Normalverteilung der nicht-dichotomen Modellvariablen (vgl. Kap. A2). Diese Annahmen gelten in der MLA separat für jede einzelne Ebene und müssen in aller Regel auch separat für jede einzelne Ebene getestet werden.
- Diese Annahmen wurden für das vorliegende Skript aus Zeitgründen nicht überprüft, sollten aber in jeder empirisch ausgerichteten MLA systematisch getestet werden (vgl. dazu Hox et al. 2018: 235ff; Urban/Mayerl 2018: 169-300).

---

<sup>29</sup> Berechnung von Pseudo-R<sup>2</sup> = (Var<sub>M1</sub> - Var<sub>M2</sub>)/(Var<sub>M1</sub>)

<sup>30</sup> Auf der home-page von Mplus (<https://statmodel.com>) wird in der Sektion “How-To” ein relativ komplexes Verfahren zur Durchführung von Chi-Quadrat-Differenzentests bei Modellen, die mit der MLR-Methodik geschätzt wurden, vorgestellt.



Tab. 21.3: Mplus-Inputfile für Modell 21:

```

TITLE:
  Model 21 (Three-level model)
  mit random intercepts, fixed slopes
  mit je zwei L1/2/3-Prädiktoren

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowses teffect
    lowses_m schses_m teff_m;

  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math ! ist die abh. Y-Variable
    female lowses ! sind L1-Prädiktoren
    teffect lowses_m ! sind L2-Prädiktoren
    schses_m teff_m ;! sind L3-Prädiktoren

  CLUSTER = schcode teachid;

  MISSING ARE . ;

  WITHIN = female lowses;
  BETWEEN = (teachid) teffect lowses_m (schcode)schses_m teff_m;
    ! in Klammern der Cluster-Identifizier

DEFINE:
  CENTER teffect teff_m schses_m lowses_m (GRANDMEAN);
    ! Zentrierung am Grand Mean
    ! anders als bei Heck/Thomas keine Zentrierung
    ! bei: lowses female

Analysis:
  TYPE = THREELEVEL;

MODEL:
  %WITHIN%
    math ON lowses female; ! für L!

  %BETWEEN teachid%
    math ON lowses_m teffect; ! für L2

  %BETWEEN schcode%
    math ON schses_m teff_m; ! für L3

OUTPUT:
  SAMPSTAT;

```

Tab. 21.4: Mplus-Ausgabe von Modellschätzung 21 (ausgewählte Bereiche):

MODEL FIT INFORMATION	
Number of Free Parameters	10
Loglikelihood	
H0 Value	-14538.451
H0 Scaling Correction Factor for MLR	1.4960
H1 Value	-14538.450
H1 Scaling Correction Factor for MLR	1.4960
Information Criteria	
Akaike (AIC)	29096.901
Bayesian (BIC)	29156.515
Sample-Size Adjusted BIC ( $n^* = (n + 2) / 24$ )	29124.741
Chi-Square Test of Model Fit	
Value	0.001*
Degrees of Freedom	0
P-Value	0.0000
Scaling Correction Factor for MLR	1.0000
* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.	
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.000
CFI/TLI	
CFI	1.000
TLI	1.000
Chi-Square Test of Model Fit for the Baseline Model	
Value	180.147
Degrees of Freedom	6
P-Value	0.0000
SRMR (Standardized Root Mean Square Residual)	
Value for Within	0.000
Value for Between Level 2	0.000
Value for Between Level 3	0.000
FORTSETZUNG —>	

Fortsetzung von Tabelle 21.4:

## MODEL RESULTS

	Estimate	Two-Tailed		P-Value
		S.E.	Est./S.E.	
Within Level				
MATH ON				
LOWSES	-13.674	1.769	-7.729	0.000
FEMALE	4.073	1.278	3.187	0.001
Residual Variances				
MATH	1419.607	74.993	18.930	0.000
Between TEACHID Level				
MATH ON				
LOWSES_M	-14.547	7.554	-1.926	0.054
TEFFECT	7.218	1.730	4.172	0.000
Residual Variances				
MATH	68.604	24.889	2.756	0.006
Between SCHCODE Level				
MATH ON				
SCHSES_M	-22.991	9.952	-2.310	0.021
TEFF_M	13.987	3.668	3.814	0.000
Intercepts				
MATH	601.865	1.635	368.097	0.000
Residual Variances				
MATH	23.171	10.719	2.162	0.031

Modell 22:

Drei-Ebenen-Modell,

Random-Intercepts- /Fixed+Random-Slopes-Modell

mit je zwei L1/2/3-Prädiktoren

(“intercepts-as-outcome model”)

L1-Gleichung  $Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + e_{ijk}$  (22.1)

L2-Gleichungen:  $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk}$  (22.2a)

$\beta_{1jk} = \gamma_{10k}$  (22.2b)

$\beta_{2jk} = \gamma_{20k}$  (22.2c)

L3-Gleichungen:  $\gamma_{00k} = \delta_{000} + \delta_{001}W1_k + \delta_{002}W2_k + v_{00k}$  (22.3a)

$\gamma_{01k} = \delta_{010} + v_{01k}$  (22.3b)

$\gamma_{02k} = \delta_{020}$  (22.3c)

$\gamma_{10k} = \delta_{100}$  (22.3d)

$\gamma_{20k} = \delta_{200}$  (22.3e)

“fixed part”:  $\delta_{000}, \delta_{001}, \delta_{002}, \delta_{010}, \delta_{020}, \delta_{100}, \delta_{200}$

“random part”:  $\text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(v_{00k}), \text{Var}(v_{01k})$

“covariance”:  $\text{Cov}(v_{00k}, v_{01k})$

(Das Modell 22 entspricht dem Modell 21, hat aber nunmehr einen Random TEFFECT-Slope und eine Kovarianz zwischen den v-Residuen.

Das Modell 22 ist gleich dem Modell 3 und dessen Formalisierung in Heck/Thomas 2015: 109-111 mit folgenden Abänderungen: a) es erfolgt ausschließlich eine Grand Mean-Zentrierung, b) die beiden dichotomen Variablen LOWSES und FEMALE werden nicht zentriert, c) es gibt eine L3-Kovarianz zwischen Intercept und TEFFECT(Z1)-Slope.)

Abb. 22.1: Graphische Skizze des Gleichungssystems von Modell 22

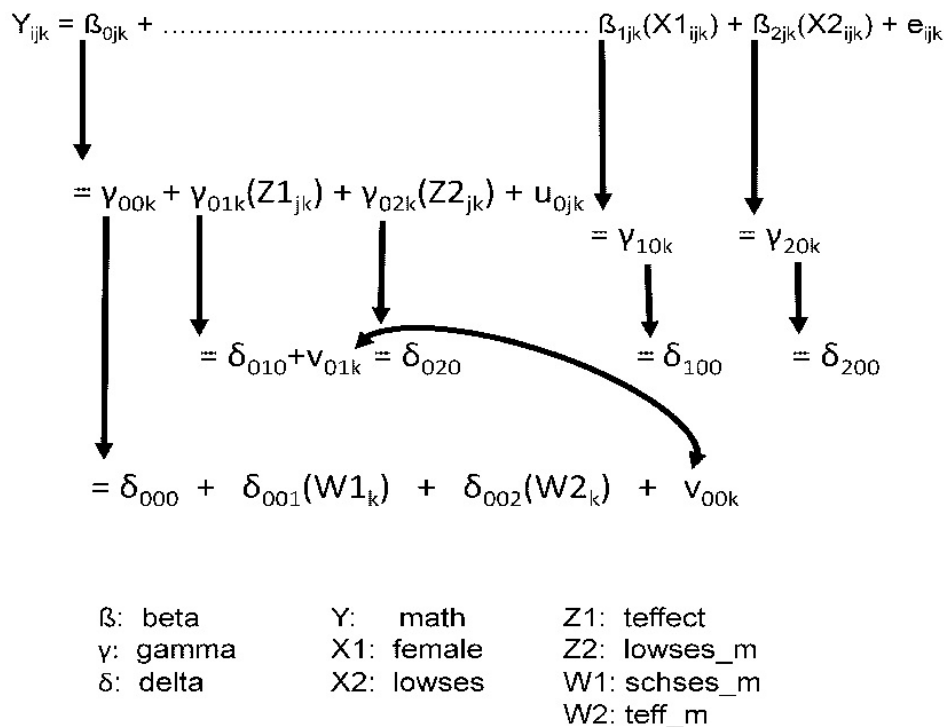
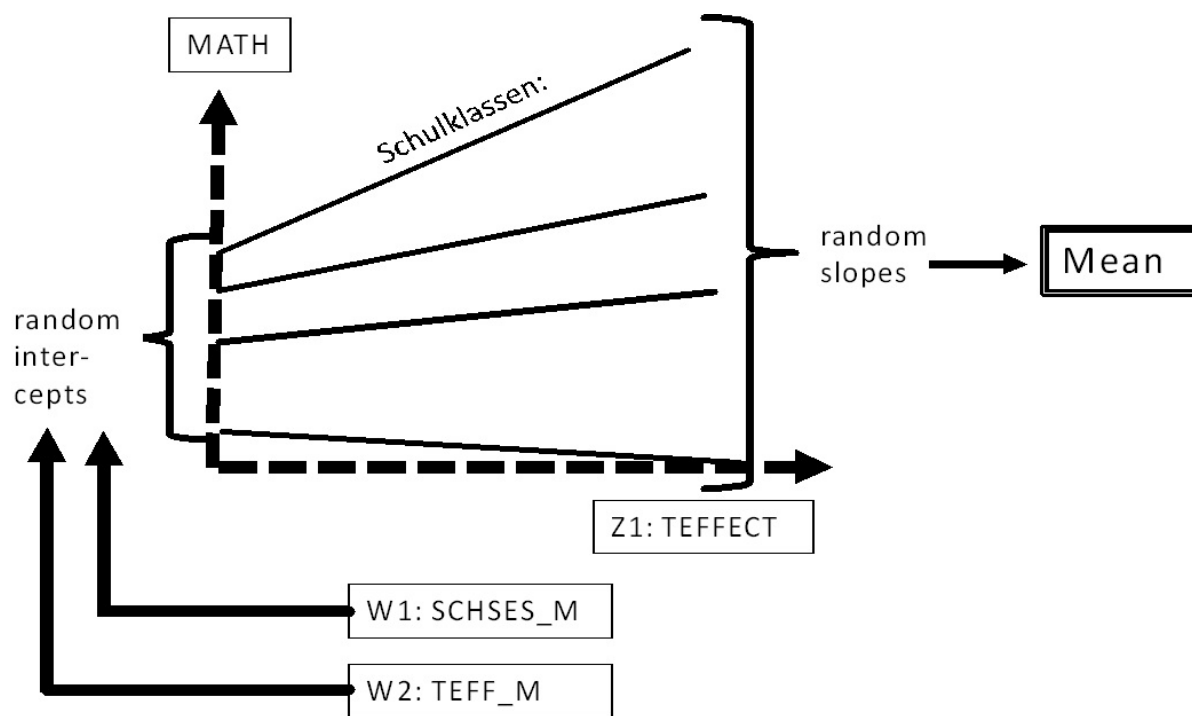


Abb. 22.2: Stilisierte Darstellung der Modell-Logik von Modell 21 (für L2/L3)

Zur Logik von Modell 22:

Das Modell 22 ist eine kleine, aber bedeutsame Weiterentwicklung von Modell 21 (vgl. dazu die Beschreibung der Modelllogik von Modell 21). Mit der Spezifikation von Modell 22 soll getestet werden, ob der Effekt von TEFFECT (Bewertung des/der Klassenlehrer:in) auf MATH (Mathematikkompetenz der einzelnen Schüler:innen) zwischen den untersuchten Schulen signifikant variiert. Dazu muss die Gl. 21.3b aus Modell 21, die den Slope von TEFFECT betrifft und diesen über alle Analyseebenen auf einen festen Wert fixiert (was in Modell 21 auch für alle anderen Slopes gilt) neu spezifiziert werden. Um die Fixierung aufzuheben, wird Gl. 21.3b um einen Random-Term erweitert, sodass der TEFFECT-Slope nunmehr zwischen den Schulen variieren kann. Vgl. dazu Gl. 22.3b (jetzt neu in Modell 22).

Mit der Ergänzung von Gl. 22.3b um einen Random-Term gibt es auf Ebene 3 nunmehr auch zwei Residual-Variablen ( $v_{00k}$ ,  $v_{01k}$ ), welche die Variation des Intercepts in Gl. 22.3a und die Variation des Slopes in Gl. 22.3b ausmachen. Wir wollen davon ausgehen, dass beide Residuen bzw. Slope und Intercept auf Ebene 3 miteinander kovariieren und deshalb als Elemente einer unstrukturierten Matrix betrachtet werden sollten.

Mit der Struktur von Modell 22 soll getestet werden, ob der TEFFECT-Slope so stark zwischen den hier untersuchten Schulen variiert, dass es sinnvoll wäre, Determinanten für diese Variation zu suchen und deren Effektstärke zu bestimmen.

Die Abbildung 22.2 verdeutlicht die Logik von Modell 22. Die Abbildung entspricht der Abbildung 21.2, zeigt aber im Koordinatensystem (gestrichelte Linien) nunmehr den L2-Zusammenhang (in Abb. 21.2 war es noch ein L1-Zusammenhang) zwischen MATH-Intercepts (random) und dem Z-Prädiktor TEFFECT. Die diesbezüglichen Steigungslinien haben "random slopes" (was neu ist).

Zur Interpretation des Gleichungssystems von Modell 22:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + e_{ijk} \quad (22.1)$$

$$\beta_{1jk} = \gamma_{10k} \quad (22.2b)$$

$$\beta_{2jk} = \gamma_{20k} \quad (22.2c)$$

$$\gamma_{10k} = \delta_{100} \quad (22.3d)$$

$$\gamma_{20k} = \delta_{200} \quad (22.3e)$$

Mit den Daten der Schüler:innen (X1, X2) wird für jede Schulklasse einer jeden Schule eine Lineargleichung zur Prognose der individuellen MATH-Testergebnisse (Y=MATH) geschätzt (Gl. 22.1). Die dichotomen Variablen X1 (FEMALE=weibl. Geschlecht) und (LOWSES = unterdurchschnittl. SES-Wert) dienen somit als Prädiktoren für MATH. Die Effektstärken dieser Prädiktoren ergeben sich aus den geschätzten Slopes. Diese Slopes gelten sowohl über alle Schulklassen als auch über alle Schulen hinweg (fixed slopes). Diese Festlegung erfolgt für X1/FEMALE durch die Gleichungen G. 22.2b und Gl. 22.3d sowie für X2/LOWSES durch die Gleichungen Gl. 22.2c und Gl. 22.3e. Das Intercept “ $\beta_0$ “ berichtet den Y-Schätzwert für die Referenzgruppe von Schüler:innen mit LOWSES=0 und FEMALE=0 sowie mit einem Durchschnittswert von 0 bei den weiteren Prädiktoren (dies ergibt sich aus der Zentrierung dieser Variablen). Die Abweichungen der geschätzten Y-Werte von den tatsächlichen Y-Werten ergeben die Residuen “e”. Daraus berechnet Mplus die ungebundene (nicht “erklärte”) Within-Varianz.

Mplus: Within Level / MATH ON LOWSES: -13.654 (p=0,000)

Mplus: Within Level / MATH ON FEMALE: 3,980 (p=0,002)

Mplus: Between SCHCODE Level / Intercepts MATH: 602,104 (p=0,000)

Mplus: Within Level / Residual Variances MATH: 1418.651 (p=0,000)

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk} \quad (22.2a)$$

$$\gamma_{01k} = \delta_{010} + v_{01k} \quad (22.3b)$$

$$\gamma_{02k} = \delta_{020} \quad (22.3c)$$

Die Intercepts aus der Linearschätzung nach Gl. 22.1 können als klassenspezifische Y=MATH-Werte (für eine bestimmte Referenzgruppe, s.o.) gelesen werden. Sie dienen in Gl. 22.2a als abh. Größe für eine zweite Linearschätzung mit den klassenspezifischen Prädiktoren Z1/TEFFECT (aggregierte Bewertung des/der Klassenlehrer:in hinsichtlich ihrer Lehrkompetenz) und Z2/LOWSES\_M (Anteil von Schüler:innen mit unterdurchschnittlichem SES-Wert). Die Slopes markieren wiederum die Effektstärken der beiden Prädiktoren. Dabei ist der Slope von LOWSES\_M fixiert und gilt für Klassen wie auch für Schulen (nach Gl. 22.3c). Der Effekt von TEFFECT ist nicht mehr fixiert, sondern ist nun (im Unterschied zu Modell 22) ein Random Effekt. Er variiert über die Schulen hinweg (nach Gl. 22.3b). Das Intercept berichtet den klas-sendurchschnittlichen  $\beta_{0jk}$ -Wert für die Referenzgruppe mit 0-Werten bei LOWSES\_M, TEFFECT, SCHSES\_M und TEFF\_M (was aufgrund der Zentrierung der vier Variablen dem jeweiligen Durchschnittswert entspricht). Da der Wert fixiert ist, ist er hier genauso groß wie in Gl. 22.1. Die aus den Residuen “u” berechnete Varianz ist die nicht “erklärte” Between-Varianz auf der L2-Ebene.

Mplus: Between TEACID Level / MATH ON LOWSES\_M: -13.691 (p=0,065)

Mplus: Between SCHCODE Level / MEANS SLOPE1: 6,330 (p=0,001)



Mplus: Between SCHCODE Level / Variances SLOPE1: 37,248 (p=0,073)

Mplus: Between TEACID Level / Residual Variances MATH: 42.669 (p=0,059)

$$\gamma_{00k} = \delta_{000} + \delta_{001} W1_k + \delta_{002} W2_k + v_{00k} \quad (22.3a)$$

Die Gleichung 22.3a ist die dritte zu schätzende Lineargleichung in Modell 22. Geschätzt wird das zu erwartende Intercept “ $\gamma_{00k}$ “ aus Gl. 22.2a. Die Prädiktoren für die Schätzung sind die schulspezifischen L3-Variablen “SCHSES\_M” (Durchschnitt der Anteile von Schüler:innen mit unterdurchschnittlichem SES-Wert pro Schule) und “TEFF\_M” (Durchschnitt der Bewertungen aller Lehrer:innen pro Schule). Die Interpretation der Gleichungskomponenten erfolgt in gleicher Weise wie oben bei Gl. 22.1 und 22.2a erläutert. Die Lineargleichung liefert auch den Schätzwert für das Intercept “ $\delta_{000}$ “, der als fixer Wert an Gl. 22.1 und Gl. 22.2a weitergegeben wird. Die aus den Residuen berechnete Varianz ist die ungebundene (nicht “erklärte”) Between-Varianz auf der L3-Ebene.

Mplus: Between SCHCODE Level / MATH ON SCHSES\_M: -25.277 (p=0,012)

Mplus: Between SCHCODE Level / MATH ON TEFF\_M: 14,770 (p=0,000)

Mplus: Between SCHCODE Level / Intercepts MATH: 602,104 (p=0,000)

Mplus: Between SCHCODE Level / Residual Variances MATH: 28.371 (p=0,015)

$$\text{Cov}(v_{00k}, v_{01k})$$

Der Term “ $\text{Cov}(v_{00k}, v_{01k})$ ” bezeichnet die Kovarianz zwischen  $v_{01k}$  und  $v_{00k}$ . Diese Kovarianz entspricht der Kovarianz zwischen dem L3-Slope (Gl. 22.3b) und dem L3-Intercept. Unterstellt wird dabei eine unstrukturierte Varianz/Kovarianz-Matrix zwischen Intercepts und Slopes.

Mplus: Between SCHCODE Level / S with MATH: -4,497 (p=0,745)



Tab. 22.1: Ergebnisse der MLA-Schätzung von Modell 22: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L3-Intercepts ( $\delta_{000}$ ) MATH	602,104	0,000
L1-Slope ( $\delta_{100}$ ) MATH ON FEMALE	3,980	0.002
L1-Slope ( $\delta_{200}$ ) MATH ON LOWSES	-13,654	0,000
L3-Slope ( $\delta_{010}$ ) Mean(Slope(TEFFECT))	6,330	0.000
L2-Slope ( $\delta_{020}$ ) MATH ON LOWSES_M	-13,691	0,065
L3-Slope ( $\delta_{001}$ ) MATH ON SCHSES_M	-25,277	0,012
L3-Slope ( $\delta_{002}$ ) MATH ON TEFF_M	14,770	0,000

Tab. 22.2: Ergebnisse der MLA-Schätzung von Modell 22: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH, $\text{Var}(e_{ijk})$	1418,651	0.000
L2-Residualvarianz, MATH, $\text{Var}(u_{ojk})$	42,669	0,059
L3-Residualvarianz, MATH, $\text{Var}(v_{00k})$	28,371	0,015
L3-Varianz: Slope(TEFFECT), $\text{Var}(v_{01k})$	37,248	0,073

Tab. 22.3: Ergebnisse der MLA-Schätzung von Modell 22: “covariances”

Kovarianzen	Schätzung	Sign.Niveau (p) zweiseitig
$\text{Cov}(v_{00k}, v_{01k})$	-4.497	0,745

Tab. 22.4: Mplus-Inputfile für Modell 22:

```

TITLE:
  Model 22 (three-level model)
  mit random intercepts, fixed+random slopes
    (wie Modell 21, aber jetzt mit random slopes für TEFFECT)
  mit je zwei L1/2/3-Prädiktoren

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsyes teffect
    lowsyes_m schses_m teff_m;

  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math ! ist die abh. Y-Variable
    female lowsyes ! sind zwei L1-Prädiktoren
    teffect lowsyes_m ! sind zwei L2-Prädiktoren
    schses_m teff_m ; ! sind zwei L3-Prädiktoren

  CLUSTER = schcode teachid;

  MISSING ARE . ;

  WITHIN = female lowsyes;
  BETWEEN = (teachid) teffect lowsyes_m (schcode)teff_m schses_m;
    ! in Klammern der Cluster-Identifizier

DEFINE:
  CENTER teffect teff_m schses_m lowsyes_m (GRANDMEAN);
    ! Zentrierung am Grand Mean
    ! anders als bei Heck/Thomas keine Zentrierung bei: lowsyes female

ANALYSIS:
  TYPE = THREELLEVEL RANDOM;
MODEL:
  %WITHIN%
    math ON lowsyes female; ! für L1
  %BETWEEN teachid%
    math ON lowsyes_m; ! für L2
    Slope | math ON teffect;
  %BETWEEN schcode%
    math ON schses_m teff_m; ! für L3
    Slope with MATH; ! Korrelation zwischen Slope(v) und L3-Intercept(v)

OUTPUT:
  SAMPSTAT;

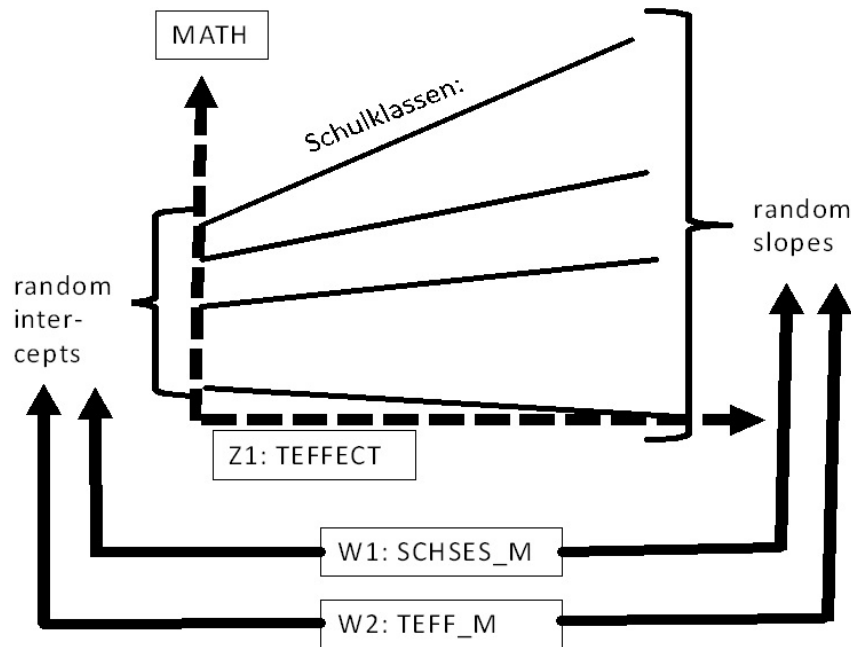
```

Tab. 22.5: Mplus-Ausgabe von Modellschätzung 22 (ausgewählte Bereiche):

MODEL FIT INFORMATION				
Number of Free Parameters				12
Loglikelihood				
H0 Value				-14535.906
H0 Scaling Correction Factor for MLR				1.3078
Information Criteria				
Akaike (AIC)				29095.811
Bayesian (BIC)				29167.348
Sample-Size Adjusted BIC				29129.219 (n* = (n + 2) / 24)
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH ON				
LOWSES	-13.654	1.769	-7.719	0.000
FEMALE	3.980	1.285	3.098	0.002
Residual Variances				
MATH	1418.651	75.126	18.883	0.000
Between TEACHID Level				
MATH ON				
LOWSES_M	-13.691	7.432	-1.842	0.065
Residual Variances				
MATH	42.669	22.621	1.886	0.059
Between SCHCODE Level				
MATH ON				
SCHSES_M	-25.277	10.007	-2.526	0.012
TEFF_M	14.770	3.600	4.102	0.000
SLOPE WITH				
MATH	-4.497	13.849	-0.325	0.745
Means				
SLOPE	6.330	1.801	3.515	0.000
Intercepts				
MATH	602.104	1.617	372.450	0.000
Variances				
SLOPE	37.248	20.760	1.794	0.073
Residual Variances				
MATH	28.371	11.641	2.437	0.015



Abb. 23.2: Stilisierte Darstellung der Modell-Logik von Modell 23 (für L2/L3)

Zur Logik von Modell 23:

Das Modell 23 unterscheidet sich allein dadurch von Modell 22, dass es die TEFFECT-Slopes als Werte einer abhängigen Variablen enthält und den Einfluss von SCHSES\_M und TEFF\_M (bzw. von W1 und W2) auf die TEFFECT-Slopes untersuchen will. Es wird also zusätzlich danach gefragt, ob der Effekt, mit dem die didaktischen Fähigkeiten des Lehrpersonals auf Klassenebene die Mathematikleistung der Schüler:innen beeinflusst (TEFFECT-Slope), vom schulbezogenen Anteil an Schüler:innen mit unterdurchschnittlichem SES (=SCHSES\_M) und/oder von der Qualität der Lehrer:innen auf Schulebene (TEFF\_M) verstärkt oder abgeschwächt wird. Dabei handelt es sich um cross-level-Effekte, weil die TEFFECT-Slopes auf L2-Ebene (= Schulklassen) liegen und die beiden Prädiktoren SCHSES\_M und TEFF\_M auf L3-Ebene (=Schulen) angesiedelt sind.

In Abbildung 23.2 werden die beiden cross-level-Effekte veranschaulicht. Die Abbildung zeigt vier Random-Slopes von TEFFECT auf MATH für vier Schulklassen, in denen die Steigungsraten unterschiedlich stark ausgeprägt sind (Linien im gestrichelt gezeichneten MATH/TEFFECT-Koordinatensystem mit unterschiedlich starken Anstiegen). Die Variation der Random-Slopes wird durch die cross-level-Effekte von SCHSES\_M und TEFF\_M geschätzt (fett gedruckte Pfeile auf der rechten Seite der Graphik). Gleichzeitig variieren auch die MATH-Intercepts auf der L2-Ebene zwischen den einzelnen Schulklassen (Y-Koordinate), sodass die Effekte von SCHSES\_M und TEFF\_M auf die Mathematikleistung der Schüler:innen ebenfalls geschätzt werden können (fett gedruckte Pfeile auf der linken Seite der Graphik).

Weitere Informationen zur Spezifikation von Modell 23 können den Erläuterungen zu Modell 21 und Modell 22 entnommen werden.

Die Abbildung 23.2 verdeutlicht die Logik von Modell 23. Insbesondere enthält sie die fett gezeichneten Pfeile, die den Einfluss von W1 und W2 auf die Intercepts von MATH und auf die Slopes von Z1/TEFFECT symbolisieren. Für weitere Erläuterungen vgl. die Informationen zu Abbildung 22.2 und Abbildung 21.2.

Zur Interpretation des Gleichungssystems von Modell 23:

Das Gleichungssystem von Modell 23 unterscheidet sich nur durch die Gl. 23.3b von den Gleichungen der Modelle 21 und 22. Dementsprechend wird hier nur die Gl. 23.3b erläutert. Zur Erläuterung der anderen Gleichungen sei auf die Ausführungen zu den Gleichungssystemen von Modell 21 und 22 verwiesen.

$$\gamma_{01k} = \delta_{010} + \delta_{011} W1_k + \delta_{012} W2_k + v_{01k} \quad (23.3b)$$

Nach dieser Lineargleichung wird die L2-Variation im Z1-Slope (TEFFECT-Slope) “ $\gamma_{01k} Z1_{jk}$ ” (vgl. Gl. 23.2a) auf die beiden L3-Lineareffekte von W1 (SCHSES\_M) und W2 (TEFF\_M) zurückgeführt. Somit handelt es sich bei den beiden Effekten um cross-level-Effekte.

Zur Interpretation der Schätzung von Modell 23:

(vgl. dazu auch die folgende Abbildung 23.3 sowie die folgenden Tabellen 23.1 und 23.2)

- 1) Die geschätzten festen (fixed) Koeffizienten entsprechen weitgehend den geschätzten fixen Koeffizienten der Modelle 21 und 22. Fast alle Interpretationshinweise zu den Resultaten der Schätzung von Modell 21 und Modell 2 sind dementsprechend auch für Modell 22 gültig.
- 2) Die Schätzwerte für die beiden oben benannten cross-level-Effekte sind:  

$$\text{Slope(TEFFECT)} = 6,58 + 6,39^{ns} \text{SCHSES\_M} \quad \text{und:}$$

$$\text{Slope(TEFFECT)} = 6,58 + 0,48^{ns} \text{TEFF\_M} \quad (\text{vgl. dazu auch Tabelle 23.1})$$

Sie sind beide eindeutig nicht-signifikant. Es ist also dementsprechend nicht davon auszugehen, dass die hier gemessenen Schul-Eigenschaften SCHSES\_M und TEFF\_M einen Einfluss auf das Klassengeschehen beim Mathematikunterricht haben (nach Gl. 23.3c).
- 3) Im Vergleich zu Modell 22 kann die Schätzung von Modell 23 die freie Varianz auf L2- und L3-Ebene nur um 2,0% bzw. 6,5% reduzieren. Das ist nicht viel. Erst recht nicht, wenn berücksichtigt wird, dass Modell 23 für seine Schätzung zwei zusätzliche freie Parameter zur Verfügung stehen. Allerdings ist die nur geringfügige Verbesserung des Schätzerfolgs sicherlich auch eine Folge der sehr schwachen und nicht-signifikanten cross-level-Effekte.
- 4) Auch ein Vergleich der AIC- und BIC-Werte von Modell 22 und Modell 23 bestätigt die Interpretation der Modellschätzung in Pkt. 3. Beide Werte nehmen im Vergleich zu denjenigen in Modell 22 sogar leicht zu (vgl. die AIC/BIC-Werte in Tab. 22.5 und in Tab.23.4)
- 5) Tabelle 23.2 zeigt, dass im Modell noch viel freie bzw. ungebundene Varianz enthalten ist. Insbesondere die Varianz der Residuen auf der L1-Ebene ist hochsignifikant. So wäre es insbesondere für diesen Varianzanteil sinnvoll, nach weiteren Prädiktoren zu suchen, die eine deutliche Erklärungsleistung auf der L1-Ebene erbringen können.
- 6) Die Interpretationshinweise zur Modellschätzung 21/Pkt.7 gelten auch für Modell 23.





Tab. 23.2: Ergebnisse der MLA-Schätzung von Modell 23: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH, Var( $e_{ijk}$ )	1419,106	0.000
L2-Residualvarianz, MATH, Var( $u_{0jk}$ )	41,769	0,056
L3-Residualvarianz, MATH, Var( $v_{00k}$ )	26,504	0,019
L3-Residualvar., Slope(TEFFECT), Var( $v_{01k}$ )	35,694	0,061

Tab. 23.3: Mplus-Inputfile für Modell 23

```

TITLE:
  Model 23 (three-level model)
  mit random intercepts/ fixed+random slopes
  (random slopes für TEFFECT-Effekt mit Prädiktoren W1 und W2
  mit je 2 L1/2/3-Prädiktoren

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsese teffect
    lowsese_m schses_m teff_m;

  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math ! ist die abh. Y-Variable
    female lowsese ! sind L1-Prädiktoren
    teffect lowsese_m ! sind L2-Prädiktoren
    schses_m teff_m ;! sind L3-Prädiktoren

  CLUSTER = schcode teachid;

  MISSING ARE . ;

  WITHIN = female lowsese;
  BETWEEN = (teachid) teffect lowsese_m (schcode)teff_m schses_m;
    ! in Klammern der Cluster-Identifizier

DEFINE:
  CENTER teffect teff_m schses_m lowsese_m (GRANDMEAN);
    ! Zentrierung am Grand Mean
    ! anders als bei Heck/Thomas keine Zentrierung bei: lowsese female

Analysis:
  TYPE = THREELLEVEL RANDOM;

MODEL:
  %WITHIN%
    math ON lowsese female; ! für L!

  %BETWEEN teachid%
    math ON lowsese_m; ! für L2
    Slope | math ON teffect;

  %BETWEEN schcode%
    math ON schses_m teff_m; ! für L3
    Slope ON schses_m teff_m ; ! neue zusätzl. Anweisung

OUTPUT:
  SAMPSTAT;

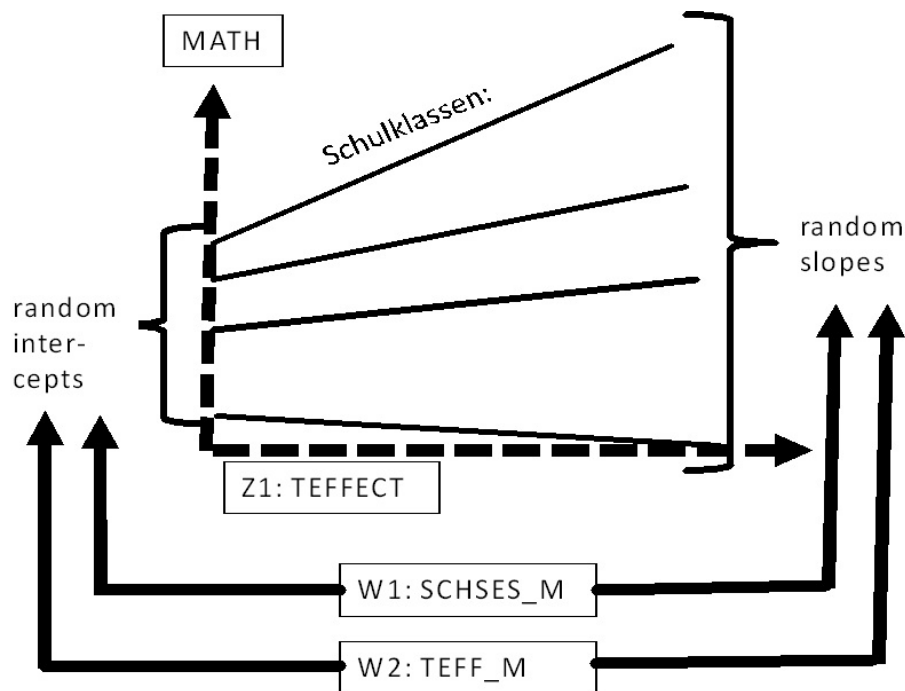
```

Tab 23.4: Mplus-Ausgabe von Modellschätzung 23 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				13
Loglikelihood				
H0 Value				-14535.082
H0 Scaling Correction Factor for MLR				1.2259
for MLR				
Information Criteria				
Akaike (AIC)				29096.165
Bayesian (BIC)				29173.663
Sample-Size Adjusted BIC				29132.357 (n* = (n + 2) / 24)
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH	ON			
LOWSES	-13.599	1.769	-7.686	0.000
FEMALE	4.038	1.286	3.139	0.002
Residual Variances				
MATH	1419.106	75.325	18.840	0.000
Between TEACHID Level				
MATH	ON			
LOWSES_M	-12.133	7.214	-1.682	0.093
Residual Variances				
MATH	41.769	21.887	1.908	0.056
Between SCHCODE Level				
SLOPE1	ON			
SCHSES_M	6.398	5.292	1.209	0.227
TEFF_M	0.485	2.743	0.177	0.860
MATH	ON			
SCHSES_M	-28.541	9.200	-3.102	0.002
TEFF_M	13.905	3.277	4.243	0.000
Intercepts				
MATH	602.240	1.607	374.788	0.000
SLOPE	6.582	1.731	3.803	0.000
Residual Variances				
MATH	26.504	11.346	2.336	0.019
SLOPE	35.694	19.035	1.875	0.061



Abb. 24.2: Stilisierte Darstellung der Modell-Logik von Modell 24 (für L2/L3)



#### Zur Logik von Modell 24:

Das Modell 24 ist identisch mit dem Modell 23. Deshalb ist auch die stilisierte Darstellung der Logik von Modell 24 (in Abb. 24.2) identisch mit derjenigen von Modell 23. Allerdings wurden nunmehr alle Z- und W-Prädiktoren neu operationalisiert. Sie wurden mittels Mediansplit dichotomisiert. Damit sind alle unabhängigen Variablen in Modell 24 dichotomisiert (die beiden X-Prädiktoren LOWSES und FEMALE sind von ihrer Konstruktionslogik her binomial skaliert). Da nunmehr alle Prädiktoren dichotomisiert sind, können ihre Effektstärken miteinander verglichen werden.

Die Abbildung 24.2 verdeutlicht die Logik von Modell 24. Sie ist identisch mit der Abbildung 23.2 (jedoch sind nunmehr alle Prädiktoren dichotomisiert)

#### Zur Interpretation des Gleichungssystems von Modell 24:

Das formale Gleichungssystem von Modell 24 ist identisch mit dem Gleichungssystem von Modell 23 und muss deshalb an dieser Stelle nicht noch einmal interpretiert werden. Die graphische Skizze des Gleichungssystems (in Abb. 24.1) ist ebenfalls identisch mit derjenigen von Modell 23.

Zur Interpretation der Schätzung von Modell 24:

(vgl. dazu auch die folgende Abbildung 24.3 sowie die folgenden Tabellen 24.1 und 24.2)

- 1) Die Schätzwerte von Modell 24 sind im Unterschied zu den Schätzwerten anderer Modelle miteinander vergleichbar, da alle Variablen mit einer 1/0-Kodierung dichotomisiert wurden.
- 2) Lehrer:innen haben einen deutlichen Einfluss auf die MATH-Fähigkeiten von Schüler:innen. Hochqualifizierte Klassenlehrer:innen heben die zu erwartende Mathe-Leistung der Schüler:innen um ca. 6 Punkte an ( $TEFFECT=6,20^{***}$ ) und auch bei einem hochqualifizierten Lehrer:innenkollegium einer Schule steigt die zu erwartende Mathe-Leistung um ca. 8 Punkte ( $TEFF\_M=8,46^{***}$ ).
- 3) Ein unterdurchschnittlicher SES verringert auf allen Analyseebenen (Schüler:innen, Klassen, Schulen) die MATH-Fähigkeiten der einzelnen Schüler:innen ( $LOWSES=-14,46^{***}$ ;  $LOWSES\_M=-10,78^*$ ;  $SCHSES\_M=-4,48^{ns}$ ), wenn auch der einzige hochsignifikante Negativ-Effekt auf der Ebene der Schüler:innen liegt: der zu erwartende MATH-Wert reduziert sich um ca. 14 Punkte, wenn ein/e Schüler:in einen unterdurchschnittlichen SES aufweist. Unabhängig vom Signifikanzniveau betrachtet, wirkt jeder SES-Prädiktor für sich alleine und wird nicht von der Effektstärke der anderen beiden SES-Prädiktoren beeinflusst. Wenn also z.B. Schüler:innen einen  $LOWSES=1$  aufweisen, wirkt  $LOWSES\_M$  (Klasse) noch insofern, als jede Steigerung des  $LOWSES\_M$ -Anteils von 0 auf 1 die zu erwartende MATH-Leistung um ca. 10 Punkte reduziert ( $LOWSES\_M=-10,78$ ).
- 4) Schülerinnen (weibl.) haben bei Konstanz aller anderen Prädiktoren einen um ca. 4 Punkte höheren MATH-Leistungsstandard als Schüler (männl.).
- 5) Auf der Ebene der Schüler:innen ist der Effekt von  $LOWSES$  mehr als dreimal so stark wie der Effekt von  $FEMALE$  ( $-14,46$  vs.  $4,02$ ), allerdings haben beide Prädiktoren unterschiedliche Wirkungsrichtungen: Ein Wert von  $LOWSES=1$  reduziert die MATH-Leistung um ca. 14 Punkte, während ein Wert von  $FEMALE=1$  die MATH-Leistung um ca. 4 Punkte erhöht.
- 6) Die AIC- und BIC-Werte von Modell 24 sind deutlich höher als die Werte von Modell 23 und signalisieren einen schlechteren Fit des Modells 24 (vgl. Kap. B5). Es scheint, als ob durch die Dichotomisierung der Variablen auch Information aufgegeben wird, was zu einer schlechteren Anpassung des Modells an die empirischen Beobachtungswerte führen muss.
- 7) Die Schätzwerte für die beiden cross-level-Effekte (vgl. dazu auch Tabelle 24.1):  
 $Slope(TEFFECT)=6,20 + 2,53^{ns}SCHSES\_M$  und:  
 $Slope(TEFFECT)=6,20 + 4,71^{ns}TEFF\_M$   
sind beide eindeutig nicht-signifikant. Es ist also dementsprechend nicht davon auszugehen, dass die hier gemessenen Schul-Eigenschaften ( $SCHSES\_M$  und  $TEFF\_M$ ) einen Einfluss auf das Klassengeschehen beim Mathematikunterricht haben (entsprechend Gl. 24.3b).

- 8) Vergleicht man die Effektstärken von allen signifikanten Effekten des Modells, so hat die soziale Herkunft der Schüler:innen den stärksten Effekt auf deren Mathematikleistung. Bei Schüler:innen mit unterdurchschnittlichem SES ist eine um ca 14 Punkte niedrigere MATH-Leistung zu erwarten als bei Schüler:innen mit überdurchschnittlichem SES.
- 9) Die Schätzungen von Modell 24 lassen sich nur bedingt mit denjenigen von Modell 23 vergleichen, weil die Skalierungen der Variablen unterschiedlich sind. Es lässt sich jedoch feststellen, dass a) alle Vorzeichen der geschätzten Parameter in beiden Modellen gleich sind, und dass b) von dreizehn, der in jedem Modell geschätzten Parameter, insgesamt fünf Parameter deutlich unterschiedliche Signifikanzen aufweisen.
- 10) Die Interpretationshinweise zur Modellschätzung 21/Pkt.7 gelten auch für Modell 24.

Abb. 24.3: Graphische Skizze des Gleichungssystems von Modell 24 mit Schätzergebnissen

$$\begin{aligned}
 Y_{ijk} &= \beta_{0jk} + \dots + \beta_{1jk}(X1_{ijk}) + \beta_{2jk}(X2_{ijk}) + e_{ijk} \\
 &= \gamma_{00k} + \gamma_{01k}(Z1_{jk}) + \gamma_{02k}(Z2_{jk}) + u_{0jk} \\
 &= \delta_{000} + \delta_{001}(W1_k) + \delta_{002}(W2_k) + v_{00k} \\
 &= \delta_{010} + \delta_{011}(W1_k) + \delta_{012}(W2_k) + v_{01k} \\
 &= \delta_{020} + \delta_{100} + \delta_{200}
 \end{aligned}$$

$\delta_{020}$  (-10,78\*)       $\delta_{100}$  (4,02\*\*\*)       $\delta_{200}$  (-14,46\*\*\*\*)

$\delta_{010}$  (6,20<sup>ns</sup>)       $\delta_{011}$  (2,53<sup>ns</sup>)       $\delta_{012}$  (4,71<sup>ns</sup>)

$\delta_{000}$  (601,51\*\*\*)       $\delta_{001}$  (-4,48<sup>ns</sup>)       $\delta_{002}$  (8,46\*)

β: beta	Y: math	Z1: teffect
γ: gamma	X1: female	Z2: lowses_m
δ: delta	X2: lowses	W1: schses_m
		W2: teff_m



Tab. 24.1: Ergebnisse der MLA-Schätzung von Modell 24: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L3-Intercepts ( $\delta_{000}$ ) MATH	601,519	0,000
L1-Slope ( $\delta_{100}$ ) MATH ON FEMALE	4,024	0,002
L1-Slope ( $\delta_{200}$ ) MATH ON LOWSES	-14,460	0,000
L3-Intercepts ( $\delta_{010}$ ) Slope(TEFFECT)	6,204	0,091
L2-Slope ( $\delta_{020}$ ) MATH ON LOWSES_M	-10,781	0,010
L3-Slope1 ( $\delta_{011}$ ) Slope1 ON SCHSES_M	2,530	0,551
L3-Slope1 ( $\delta_{012}$ ) Slope1 ON TEFF_M	4,712	0,338
L3-Slope ( $\delta_{001}$ ) MATH ON SCHSES_M	-4,480	0,277
L3-Slope ( $\delta_{002}$ ) MATH ON TEFF_M	8,460	0,039

Tab. 24.2: Ergebnisse der MLA-Schätzung von Modell 24: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH, $\text{Var}(e_{ijk})$	1418,047	0,000
L2-Residualvarianz, MATH, $\text{Var}(u_{0jk})$	90,814	0,008
L3-Residualvarianz, MATH, $\text{Var}(v_{00k})$	47,916	0,020
L3-Residualvarianz, Slope(TEFFECT) $\text{Var}(v_{01k})$	0,872	0,981

Tab. 24.3: Mplus-Inputfile für Modell 24:

```

TITLE:
  Model 24 (three-level model)
  mit random intercepts / fixed+random slopes
  (Random Slopes für TEFFECT-Effekt, Prädiktoren W1 und W2)
  mit je zwei L1/2/3-Prädiktoren (alle dichotomisiert, teilw. Mediansplit)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsese teffect
    lowsese_m schses_m teff_m;

  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math ! ist die abh. Y-Variable
    female lowsese ! sind L1-Prädiktoren
    z1 z2 w1 w2 ; ! werden alle über Mediansplit hergestellt (s.u.)

  CLUSTER = schcode teachid;

  MISSING ARE . ;

  WITHIN = female lowsese;
  BETWEEN = (teachid) z1 z2 (schcode)w1 w2;
    ! in Klammern ist der Cluster-Identifizier

DEFINE:
  ! jetzt Dichotomisierung über Mediansplit:
  IF(teffect LE -.09)THEN z1 =0;
  IF(teffect GT -.09)THEN z1 =1;
  IF(lowsese_m LE 0.33) THEN z2 =0;
  IF(lowsese_m GT 0.33) THEN z2 =1;
  IF(schses_m LE 0.36)THEN w1 =0;
  IF(schses_m GT 0.36)THEN w1 =1;
  IF(teff_m LE 0.01) THEN w2 = 0;
  IF(teff_m GT 0.01) THEN w2 = 1;

ANALYSIS:
  TYPE = THREELLEVEL RANDOM;

MODEL:
  %WITHIN%
    math ON lowsese female; ! für L1

  %BETWEEN teachid%
    math ON z2; ! für L2
    Slope | math ON z1;

```

FORTSETZUNG -&gt;

Fortsetzung von Tabelle 24.3:

%BETWEEN schcode%

math ON w1 w2; ! für L3

Slope ON w1 w2 ; ! neue zusätzl. Anweisung

OUTPUT:

SAMPSTAT;

Tab. 24.4: Mplus-Ausgabe von Modellschätzung 24 (ausgewählte Bereiche):

MODEL FIT INFORMATION				
Number of Free Parameters				13
Loglikelihood				
H0 Value				-14554.679
H0 Scaling Correction Factor for MLR				1.2763
Information Criteria				
Akaike (AIC)				29135.359
Bayesian (BIC)				29212.856
Sample-Size Adjusted BIC				29171.551
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH	ON			
LOWSES	-14.460	1.777	-8.136	0.000
FEMALE	4.024	1.279	3.146	0.002
Residual Variances				
MATH	1418.047	76.121	18.629	0.000
Between TEACHID Level				
MATH	ON			
Z2	-10.781	4.205	-2.564	0.010
Residual Variances				
MATH	90.814	34.282	2.649	0.008
Between SCHCODE Level				
SLOPE	ON			
W1	2.530	4.245	0.596	0.551
W2	4.712	4.922	0.957	0.338
MATH	ON			
W1	-4.480	4.123	-1.086	0.277
W2	8.460	4.090	2.068	0.039
Intercepts				
MATH	601.519	2.557	235.229	0.000
SLOPE	6.204	3.676	1.688	0.091
Residual Variances				
MATH	47.916	20.607	2.325	0.020
SLOPE	0.872	36.327	0.024	0.981

Modell 25:

Drei-Ebenen-Modell,

Random-Intercepts-/Fixed+Random-Slopes-Modell

mit je zwei L1/2/3-Prädiktoren (dichotom)

(Slope “ $\beta_{2jk}$ ” variiert auf L2 und L3, ohne Prädiktoren)

(“intercepts-as-outcome model”)

L1-Gleichung  $Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + e_{ijk}$  (25.1)

L2-Gleichungen:  $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk}$  (25.2a)

$\beta_{1jk} = \gamma_{10k}$  (25.2b)

$\beta_{2jk} = \gamma_{20k} + u_{2jk}$  (25.2c)

L3-Gleichungen:  $\gamma_{00k} = \delta_{000} + \delta_{001}W1_k + \delta_{002}W2_k + v_{00k}$  (25.3a)

$\gamma_{01k} = \delta_{010}$  (25.3b)

$\gamma_{02k} = \delta_{020}$  (25.3c)

$\gamma_{10k} = \delta_{100}$  (25.3d)

$\gamma_{20k} = \delta_{200} + v_{20k}$  (25.3e)

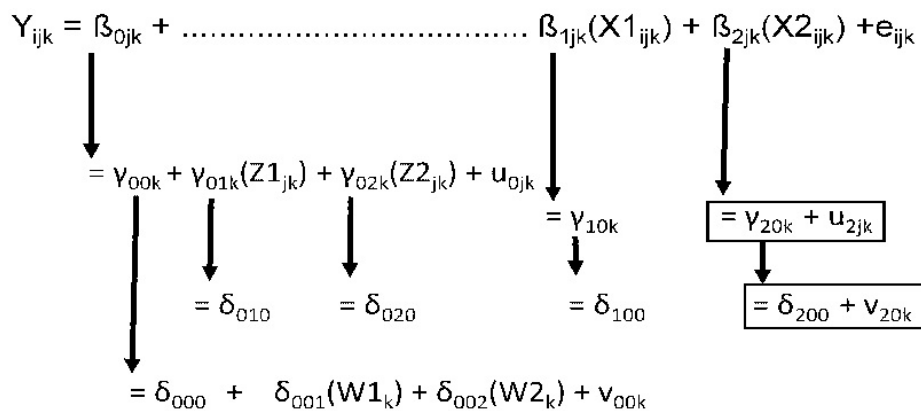
“fixed part”:  $\delta_{000}, \delta_{001}, \delta_{002}, \delta_{010}, \delta_{020}, \delta_{100}, \delta_{200}$

“random part”:  $\text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(u_{2jk}), \text{Var}(u_{20k}), \text{Var}(v_{00k})$

(Das Modell 25 ist analog zu Modell 23 konstruiert. Auch in Modell 25 sind alle Prädiktoren dichotomisiert (ohne Zentrierung). Allerdings ist jetzt der Slope  $\beta_{2jk}$  nicht mehr fixiert und variiert auf den Ebenen L2 und L3.

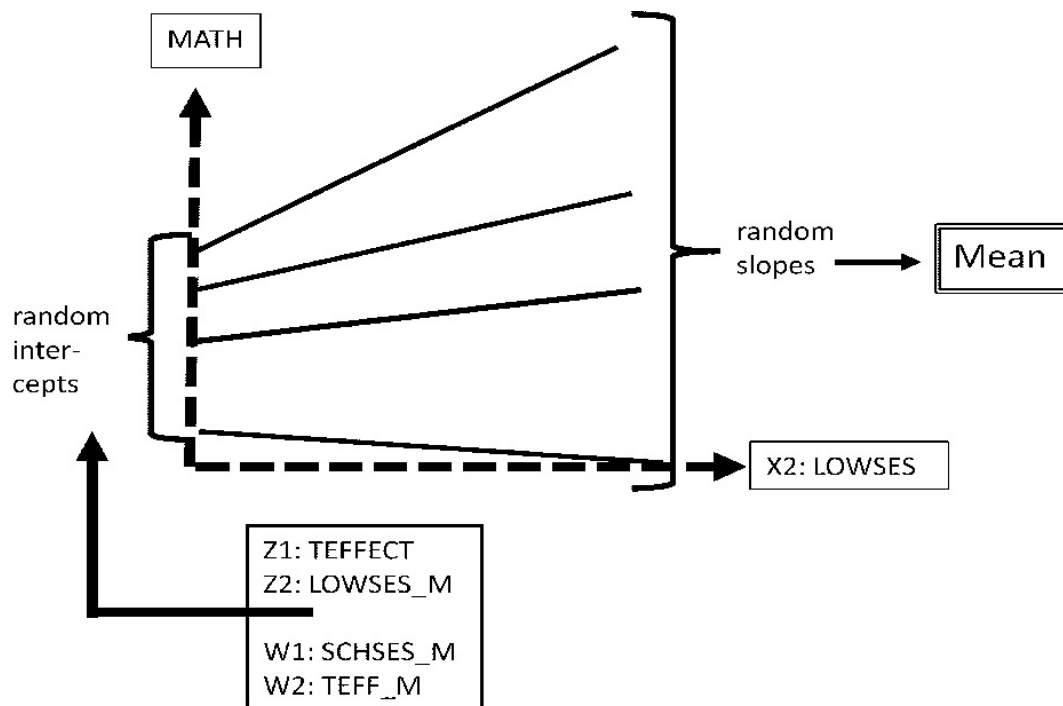
Das Modell 25 entspricht dem Modell 2 und dessen Formalisierung in Heck/Thomas 2015: 105-107 mit folgenden Abänderungen: a) alle Variablen sind dichotomisiert und nicht zentriert, b) der Slope  $\beta_{2jk}$  ist nicht mehr fixiert und variiert auf den Ebenen L2 und L3.)

Abb. 25.1: Graphische Skizze des Gleichungssystems von Modell 25



$\beta$ : beta	$\gamma$ : gamma	$\delta$ : delta	$Y$ : math	$X1$ : female	$X2$ : lows	$Z1$ : teffect	$Z2$ : lows_m	$W1$ : schses_m	$W2$ : teff_m
----------------	------------------	------------------	------------	---------------	-------------	----------------	---------------	-----------------	---------------

Abb. 25.2: Stilisierte Darstellung der Modell-Logik von Modell 25



### Zur Logik von Modell 25:

Mit Modell 25 und dem folgenden Modell 26 soll die Hypothese getestet werden, dass der Effekt, der von einer sozialen Herkunft mit unterdurchschnittlichem SES ausgeht und die MATH-Leistungen von Schüler:innen negativ beeinflusst, zusätzlich verstärkt werden kann, wenn der Anteil von Schüler:innen mit unterdurchschnittlichem SES in einer Schulklasse besonders hoch ist. Eine solche Verstärkung könnte dadurch ausgelöst werden, dass ein hoher Anteil von Schüler:innen aus Familien mit unterdurchschnittlichem SES das Lernklima innerhalb einer Schulklasse negativ beeinflusst und somit die Motivation zu besonderen Anstrengungen im Mathematik-Unterricht eher reduziert.

Um diese Hypothese zu testen, soll mit Modell 25 zunächst untersucht werden, ob die Varianz der Slopes von LOWSES auf den Ebenen von L2 und L3 genügend groß ist, damit die LOWSES-Slopes als abh. Variablen, die von Prädiktoren beeinflusst werden, analysiert werden können.

Die Abbildung 25.2 verdeutlicht die Logik von Modell 25. Von der Konstruktion her ist Abbildung 25.2 ähnlich den Abbildungen 24.2 und 23.2. Sie unterscheidet sich dadurch, dass die Steigungslinien innerhalb des Koordinatensystems die Effekte (random) von X2/LOWSES auf MATH (random) und damit auch deren Varianz symbolisieren. Jede Schulklasse weist eine eigene Steigungslinie auf, sodass der Mean aller Slopes den durchschnittlichen Effekt von LOWSES auf MATH markiert.

Zur Interpretation des Gleichungssystems von Modell 25:

$$\beta_{2jk} = \gamma_{20k} + u_{2jk} \quad (25.2c)$$

$$\gamma_{20k} = \delta_{200} + v_{20k} \quad (25.3e)$$

Um zu testen, ob die Varianz des Slopes von LOWSES genügend groß ist, muss die Fixierung von  $\beta_{2jk}$  aufgehoben werden. Das geschieht dadurch, dass mit Gl. 25.2c und Gl. 25.3e der LOWSES-Slope um zwei Random-Terme erweitert wird (vgl. die Rechtecke in Abb. 25.1) und somit zwischen Schulklassen und Schulen variieren kann. Die Schätzung der freien Varianzen erbringt jedoch enttäuschende Ergebnisse. Die beiden Varianzen sind deutlich nicht signifikant ( $p=0,977$  bzw.  $0,984$ ). Nur der geschätzte Mittelwert aller LOWSES-Slopes ist mit einem Wert von  $-13,65$  hochsignifikant.

Mplus: Between TEACHID Level / Variances S1: 1,124 ( $p=0,977$ )

Mplus: Between SCHCODE Level / Variances S1: 0,781 ( $p=0,984$ )

Mplus: Between SCHCODE Level / MEANS S1:  $-13,657$  ( $p=0,000$ )

Ansonsten ist das formale Gleichungssystem von Modell 25 (bis auf die oben erläuterten Modifikationen) identisch mit dem Gleichungssystem von Modell 21 und muss deshalb an dieser Stelle nicht noch einmal erörtert werden.

Zur Interpretation der Schätzung von Modell 25:

(vgl. dazu auch die folgende Abbildung 25.3 sowie die folgenden Tabellen 25.1 und 25.2)

- 1) Bezüglich unserer Ausgangsfrage variiert der Slope " $b_{2jk}$ " weder zwischen den einzelnen Schulklassen noch zwischen den einzelnen Schulen in signifikanter Weise. Die Irrtumswahrscheinlichkeiten von  $0,977$  und  $0,984$  (vgl. Tabelle 25.2) verweisen darauf, dass sich die entsprechenden Varianzen kaum von  $0,00$  unterscheiden. Allerdings ist der Effekt von LOWSES auf MATH " $\beta_{2jk}$ " bedeutsam und auch hochgradig signifikant: der zu erwartende MATH-Wert reduziert sich über alle Slopes im Mittel um  $-13,65$ , wenn Schüler:innen aus einer Familie mit unterdurchschnittlichem SES kommen (vgl. Tabelle 25.1).
- 2) Obwohl es aufgrund der unter Pkt. (1) erläuterten Schätzergebnisse nicht sinnvoll ist, die zuvor aufgestellte Hypothese zur Verstärkung des negativen LOWSES-Effekts statistisch zu überprüfen, wollen wir mit dem folgenden Modell 26 dennoch aufzeigen, in welcher Weise ein entsprechender MLA-Test im Prinzip durchgeführt werden kann.
- 3) Ein Vergleich der AIC- und BIC-Werte von Modell 24 und Modell 25 zeigt, dass beide Modelle einen vergleichbar guten Fit aufweisen (vgl. Kap. B5).
- 4) Zu weiteren Erläuterungen der Schätzung von Modell 25 vgl. die Hinweise zu der Schätzung von Modell 21.





Tab. 25.2: Ergebnisse der MLA-Schätzung von Modell 25: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH Var( $e_{ijk}$ )	1417,340	0.000
L2-Residualvarianz, MATH, Var( $u_{0jk}$ )	93,180	0,005
L3-Residualvarianz, MATH, Var( $v_{00k}$ )	47,109	0,038
L2-Slopevarianz, LOWSES Var( $u_{2jk}$ )	1,124	0,977
L3-Slopevarianz, LOWSES Var( $v_{20k}$ )	0,781	0,984

Tab. 25.3: Mplus-Inputfile für Modell 25:

```

TITLE:
  Model 25 (three-level model)
  mit random intercepts / fixed+random slopes (random slopes für beta(lowses) )
  mit je zwei L1/2/3 Prädiktoren (alle dichotomisiert: Mediansplit bei L2/L3-Prädiktoren)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;
VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsyes teffect
    lowsyes_m schses_m teff_m;

  USEVARIABLES ARE
    teachid      ! ist die L2-Clustervariable
    schcode      ! ist die L3-Clustervariable
    math         ! ist die abh. Y-Variable
    female lowsyes ! sind L1-Prädiktoren
    z1 z2 w1 w2 ; ! werden alle über Mediansplit hergestellt

  CLUSTER = schcode teachid;
  MISSING ARE . ;

  WITHIN = female lowsyes;
  BETWEEN = (teachid) z1 z2 (schcode) w1 w2; ! in Klammern der Cluster-Identifizier

DEFINE: ! jetzt Dichotomisierung über Mediansplit:
  IF(teffect LE -.09)THEN z1 =0;
  IF(teffect GT -.09)THEN z1 =1;
  IF(lowsyes_m LE 0.33) THEN z2 =0;
  IF(lowsyes_m GT 0.33) THEN z2 =1;
  IF(schses_m LE 0.36)THEN w1 =0;
  IF(schses_m GT 0.36)THEN w1 =1;
  IF(teff_m LE 0.01) THEN w2 = 0;
  IF(teff_m GT 0.01) THEN w2 = 1;

ANALYSIS:
  TYPE = THREELEVEL RANDOM;
MODEL:
  %WITHIN%
    math ON female; ! für L1
    S1 | math ON lowsyes;

  %BETWEEN teachid%
    math ON z1 z2; ! für L2

  %BETWEEN schcode%
    math ON w1 w2; ! für L3

OUTPUT:
  SAMPSTAT;

```

Tab 25.4: Mplus-Ausgabe von Modellschätzung 25 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				12
Loglikelihood				
H0 Value				-14555.115
H0 Scaling Correction Factor for MLR				1.3807
Information Criteria				
Akaike (AIC)				29134.230
Bayesian (BIC)				29205.767
Sample-Size Adjusted BIC				29167.638 (n* = (n + 2) / 24)
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH ON				
FEMALE	4.021	1.268	3.172	0.002
Residual Variances				
MATH	1417.340	74.309	19.074	0.000
Between TEACHID Level				
MATH ON				
Z1	11.171	2.566	4.354	0.000
Z2	-10.683	4.220	-2.531	0.011
Variances				
S1	1.124	39.538	0.028	0.977
Residual Variances				
MATH	93.180	32.838	2.838	0.005
Between SCHCODE Level				
MATH ON				
W1	-3.849	3.796	-1.014	0.311
W2	10.195	3.027	3.368	0.001
Means				
S1	-13.657	1.771	-7.711	0.000
Intercepts				
MATH	599.663	2.402	249.611	0.000
Variances				
S1	0.781	38.962	0.020	0.984
Residual Variances				
MATH	47.109	22.746	2.071	0.038

Modell 26:

Drei-Ebenen-Modell,

Random-Intercepts-/Fixed+Random-Slopes Modell

mit je zwei L1/2/3-Prädiktoren (dichotom)

(“intercepts-and-slopes-as-outcomes model”

(“cross-level-effect”, Prädiktor: Z2)

L1-Gleichung:  $Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + e_{ijk}$  (26.1)

L2-Gleichungen:  $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk}$  (26.2a)

$\beta_{1jk} = \gamma_{10k}$  (26.2b)

$\beta_{2jk} = \gamma_{20k} + \gamma_{30k}Z2_{jk} + u_{3jk}$  (26.2c)

L3-Gleichungen:  $\gamma_{00k} = \delta_{000} + \delta_{001}W1_k + \delta_{002}W2_k + v_{00k}$  (26.3a)

$\gamma_{01k} = \delta_{010}$  (26.3b)

$\gamma_{02k} = \delta_{020}$  (26.3c)

$\gamma_{10k} = \delta_{100}$  (26.3d)

$\gamma_{20k} = \delta_{200}$  (26.3e)

$\gamma_{30k} = \delta_{300} + v_{30k}$  (26.3f)

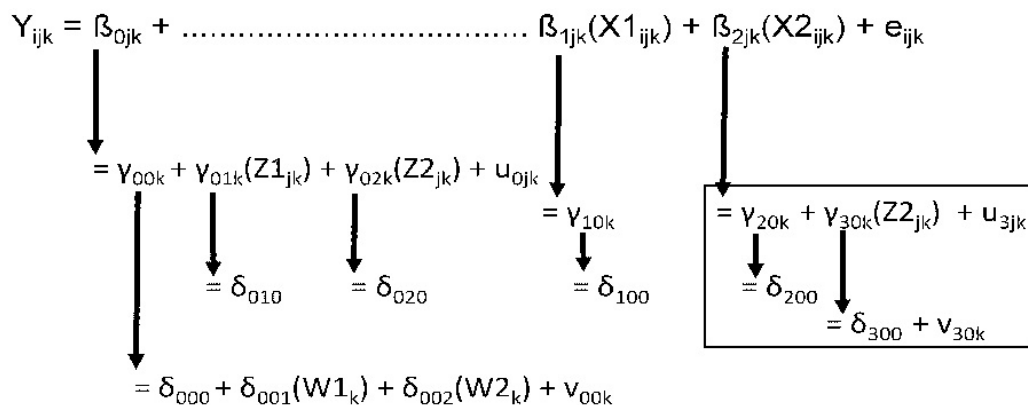
“fixed part”:  $\delta_{000}, \delta_{001}, \delta_{002}, \delta_{010}, \delta_{020}, \delta_{100}, \delta_{200}, \delta_{300}$

“random part”:  $Var(e_{ijk}), Var(u_{0jk}), Var(u_{3jk}), Var(v_{00k}), Var(v_{30k})$

(Das Modell 26 entspricht dem Modell 25 mit der Modifikation, dass der LOWSES-Slope nunmehr dem cross-level-effect von LOWSES\_M unterliegt.

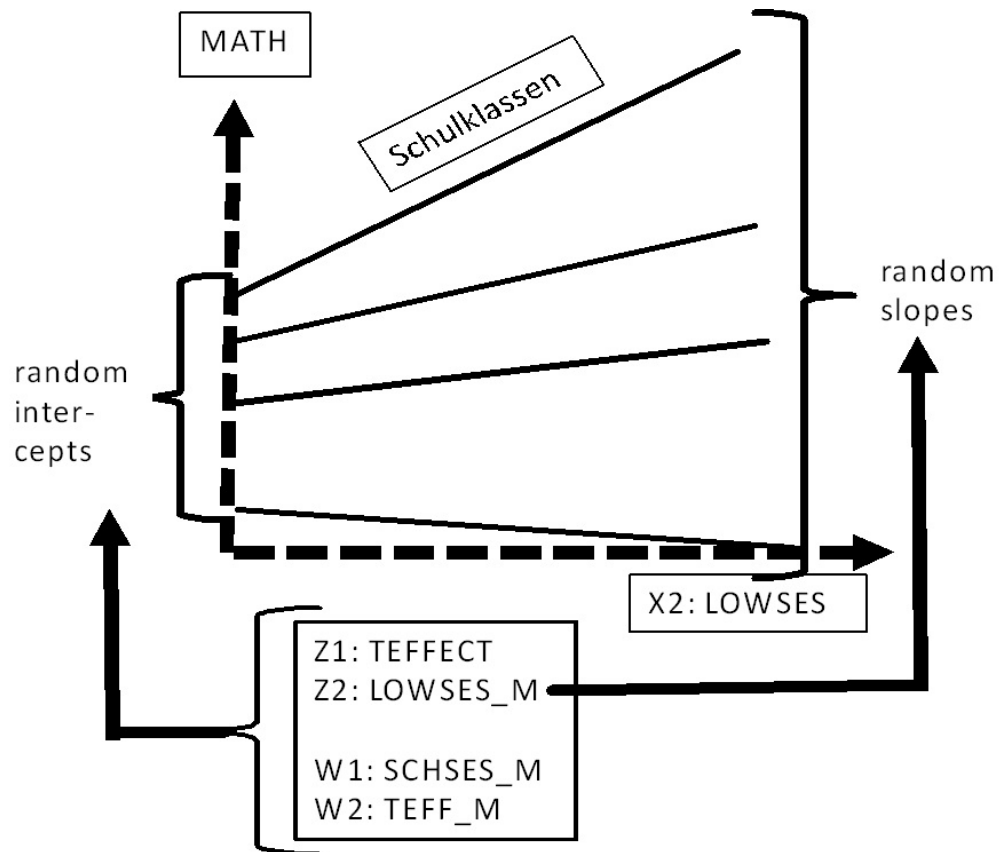
Das Modell 26 orientiert sich an dem Modell 2 und dessen Formalisierung in Heck/Thomas 2015: 105-107 mit folgenden Abänderungen: a) alle Variablen sind dichotomisiert und nicht zentriert, b) der Slope  $\beta_{2jk}$  ist nicht mehr fixiert und variiert auf den Ebenen L2 und L3.)

Abb. 26.1: Graphische Skizze des Gleichungssystems von Modell 26



$\beta$ : beta	Y: math	Z1: teffect
$\gamma$ : gamma	X1: female	Z2: lowses_m
$\delta$ : delta	X2: lowses	W1: schses_m
		W2: teff_m

Abb. 26.2: Stilisierte Darstellung der Modell-Logik von Modell 26



#### Zur Logik von Modell 26:

Nachdem mit der Modellschätzung 25 nachgewiesen wurde, dass die Varianz vom Slope  $b_{2jk}$  eigentlich nicht ausreichend ist, um den Einfluss von Prädiktoren auf den LOWSES-Slope zu testen, wir aber beschlossen haben, es dennoch zu wagen, um zumindest die dementsprechende Technik hier beschreiben zu können, wollen wir diese Technik im Folgenden auch vorführen. Es geht also darum zu untersuchen, ob der LOWSES-Slope bzw. der Effekt der Zugehörigkeit zu einer unterdurchschnittlichen Klassenlage (der die Schulleistung im Fach "Mathematik" beeinflusst) dadurch verstärkt oder abgeschwächt werden kann, dass die jeweilige Schulklasse einen hohen Anteil von Schüler:innen aus Familien mit unterdurchschnittlichem SES aufweist.

Die Abbildung 26.2 verdeutlicht die Logik von Modell 26. Sie zeigt wiederum das X/Y-Koordinatensystem (gestrichelte Linien) mit den Steigungslinien (durchgezogene schmale Linien) für den Effekt von X2/LOWSES auf MATH (für die Schüler einer jeden Schulklasse jeweils eine Steigungslinie bzw. einen Slope). Die Slopes haben dabei eine unterschiedlich starke Steigung (random slopes). Von Interesse ist in Abb. 26.2 vor allem, ob der Prädiktor Z2/LOWSES M einen Effekt auf das Ausmaß der Steigungen hat (breit gezeichnete Linie mit Pfeilspitze auf der rechten Seite der Abbildung).

Zur Interpretation des Gleichungssystems von Modell 26:

$$\beta_{2jk} = \gamma_{20k} + \gamma_{30k} Z2_{jk} + u_{3jk} \quad (26.2c)$$

$$\gamma_{30k} = \delta_{300} + v_{30k} \quad (26.3f)$$

Das Gleichungssystem zur Schätzung von Modell 26 enthält im Vergleich zu Modell 25 zwei neue Gleichungen (Gl. 26.2c und Gl. 26.3f), die auch in der graphischen Skizze zu Modell 26 im dort neu eingezeichneten Rechteck zu finden sind. Die Gleichung 26.2c legt fest, dass der LOWSES-Slope vom Prädiktor Z2 mit einer Stärke von  $\gamma_{30k}$  auf der Klassenebene beeinflusst werden kann und auch über die verschiedenen Klassenverbände hinweg variiert. Zusätzlich bestimmt Gleichung 26.3f, dass dieser Einfluss zwischen den einzelnen Schulen mit einer Varianz von  $\text{Var}(v_{30k})$  variiert.

Mplus: Between TEACHID Level / SLOPE ON Z2: -3,027 (p=0,424)

Mplus: Between TEACHID Level / Residual Variances SLOPE: 1,105 (p=0,978)

Mplus: Between SCHCODE Level / Variances SLOPE: 0,841 (p=0,982)

Ansonsten ist das formale Gleichungssystem von Modell 26 (bis auf die oben erläuterten Modifikationen) identisch mit dem Gleichungssystem von Modell 21 und muss deshalb an dieser Stelle nicht noch einmal interpretiert werden.

Zur Interpretation der Schätzung von Modell 26:

(vgl. dazu auch die folgende Abbildung 26.3 sowie die folgenden Tabellen 26.1 und 26.2)

- 1) Wie nicht anders zu erwarten war, ist der Z2-Effekt auf den LOWSES-Slope substanziell und statistisch bedeutungslos. Er hat zwar das "richtige" Vorzeichen, nach dem bei einem hohen Anteil von LOWSES-Mitschüler:innen die Abschwächung der zu prognostizierenden Mathematik-Leistung bei Schüler:innen mit Herkunft aus LOWSES-Familien noch einmal um -3,02 Punkte sinkt (von -11,78 auf -14,80 Punkte), jedoch ist das Ausmaß der negativen Verstärkung sehr deutlich nicht-signifikant (p=0,424).
- 2) Auch die Signifikanzen der Varianz des LOWESE\_M-Effekts sind auf Klassenebene und auf Schulebene sehr deutlich nicht-signifikant (p=0,978 bzw. p=0,982).
- 3) Unabhängig vom Anteil von LOWSES-Mitschüler:innen sinkt die Mathematik-Leistung um 11,78 Punkte, wenn Schüler:innen aus Haushalten mit niedrigem SES kommen.
- 4) Ein Vergleich der AIC- und BIC-Werte von Modell 25 und Modell 26 zeigt, dass beide Modelle einen vergleichbar guten Fit aufweisen (vgl. Kap. B5).
- 5) Zu weiteren Erläuterungen der Schätzung von Modell 26 vgl. die Hinweise zu der Schätzung von Modell 21 (dort: Pkt. 7).



Tab. 26.2: Ergebnisse der MLA-Schätzung von Modell 26: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH, $\text{Var}(e_{ijk})$	1417,542	0.000
L2-Residualvarianz, MATH, $\text{Var}(u_{0jk})$	93,338	0,004
L3-Residualvarianz, MATH, $\text{Var}(v_{00k})$	47,069	0,035
L2-Residualvarianz, Slope(LOWSES), $\text{Var}(u_{3jk})$	1,105	0,978
L3-Slopevarianz, LOWSES_M $\text{Var}(v_{30k})$	0,841	0,982



Tab. 26.3: Mplus-Inputfile für Modell 26:

```

TITLE:
  Model 26
  mit random intercepts/ fixed+random slopes
  (cross-level-effect: LOWSES-Slope mit Prädiktor LOWSES_M)
  mit je zwei L1/2/3-Prädiktoren (alle Variablen/Prädiktoren dichotomisiert)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsese teffect
    lowsese_m schses_m teff_m;

  USEVARIABLES ARE
    teachid ! ist die L2-Clustervariable
    schcode ! ist die L3-Clustervariable
    math ! ist die abh. Y-Variable
    female lowsese ! sind L1-Prädiktoren, dichotom
    z1 z2 w1 w2 ; ! werden alle über Mediansplit hergestellt (s.u.

  CLUSTER = schcode teachid;
  MISSING ARE . ;

  WITHIN = female lowsese;
  BETWEEN = (teachid) z1 z2 (schcode)w1 w2; ! in Klammern der Cluster-Identifizier
DEFINE: ! jetzt Dichotomisierung über Mediansplit:
  IF(teffect LE -.09)THEN z1 =0;
  IF(teffect GT -.09)THEN z1 =1;
  IF(lowsese_m LE 0.33) THEN z2 =0;
  IF(lowsese_m GT 0.33) THEN z2 =1;
  IF(schses_m LE 0.36)THEN w1 =0;
  IF(schses_m GT 0.36)THEN w1 =1;
  IF(teff_m LE 0.01) THEN w2 = 0;
  IF(teff_m GT 0.01) THEN w2 = 1;

ANALYSIS:
  TYPE = THREELLEVEL RANDOM;

MODEL:
  %WITHIN%
    math ON female; ! für L1
    Slope | math ON lowsese;
  %BETWEEN teachid%
    math ON z2 z1; ! für L2
    Slope ON z2;
  %BETWEEN schcode%
    math ON w1 w2; ! für L3

OUTPUT:
  SAMPSTAT;

```

Tab. 26.4: Mplus-Ausgabe von Modellschätzung 26 (ausgewählte Bereiche):

MODEL FIT INFORMATION				
Number of Free Parameters				13
Loglikelihood				
H0 Value				-14555.381
H0 Scaling Correction Factor for MLR				1.3565
Information Criteria				
Akaike (AIC)				29136.763
Bayesian (BIC)				29214.260
Sample-Size Adjusted BIC				29172.955
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH	ON			
FEMALE	3.999	1.277	3.132	0.002
Residual Variances				
MATH	1417.542	74.281	19.083	0.000
Between TEACHID Level				
SLOPE	ON			
Z2	-3.027	3.784	-0.800	0.424
MATH	ON			
Z2	-9.694	4.331	-2.239	0.025
Z1	11.161	2.569	4.344	0.000
Residual Variances				
MATH	93.338	32.632	2.860	0.004
SLOPE	1.105	39.396	0.028	0.978
Between SCHCODE Level				
MATH	ON			
W1	-3.828	3.794	-1.009	0.313
W2	10.193	3.038	3.355	0.001
Means				
SLOPE	-11.784	3.140	-3.754	0.000
Intercepts				
MATH	599.380	2.420	247.705	0.000
Variances				
SLOPE	0.841	37.465	0.022	0.982
Residual Variances				
MATH	47.069	22.373	2.104	0.035

Modell 27:

Drei-Ebenen-Modell,

Random-Intercepts/Fixed Slopes Modell,

mit je zwei L1/2/3 Prädiktoren (dichotom)

(“intercepts-as-outcomes”)

(mit Interaktionsvariable “X1\*X2”)

L1-Gleichung:  $Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + \beta_{3jk}X1X2_{ijk} + e_{ijk}$  (27.1)

L2-Gleichungen:  $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk}$  (27.2a)

$\beta_{1jk} = \gamma_{10k}$  (27.2b)

$\beta_{2jk} = \gamma_{20k}$  (27.2c)

$\beta_{3jk} = \gamma_{30k}$  (27.2d)

L3-Gleichungen:  $\gamma_{00k} = \delta_{000} + \delta_{001}W1_k + \delta_{002}W2_k + v_{00k}$  (27.3a)

$\gamma_{01k} = \delta_{010}$  (27.3b)

$\gamma_{02k} = \delta_{020}$  (27.3c)

$\gamma_{10k} = \delta_{100}$  (27.3d)

$\gamma_{20k} = \delta_{200}$  (27.3e)

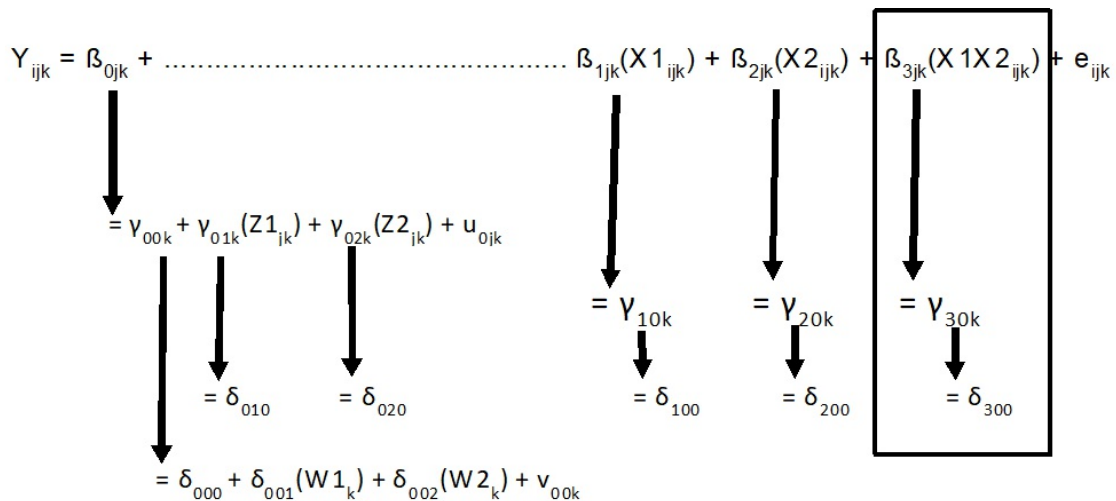
$\gamma_{30k} = \delta_{300}$  (27.3f)

“fixed part”:  $\delta_{000}, \delta_{001}, \delta_{002}, \delta_{010}, \delta_{020}, \delta_{100}, \delta_{200}, \delta_{300}$

“random part”:  $\text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(v_{00k})$

(Das Modell 27 entspricht dem Modell 21 mit der Modifikation, dass nunmehr auch der Interaktionseffekt “X1\*X2“ (=MALE\*LOWSES) hinzugekommen ist. Das Modell 27 orientiert sich an dem Modell 2 und dessen Formalisierung in Heck/Thomas 2015: 105-107, allerdings sind in Modell 27 nunmehr alle Prädiktoren dichotomisiert und auch nicht zentriert)

Abb. 27.1: Graphische Skizze des Gleichungssystems von Modell 27



$\beta$ : beta	Y: math	Z1: teffect
$\gamma$ : gamma	X1: male	Z2: lowsyes_m
$\delta$ : delta	X2: lowsyes	W1: schses_m
	X1X2: male×lowsyes	W2: teff_m



des Koordinatensystems symbolisiert. Alle Steigungslinien haben die gleiche Steigungsrate (fixed slopes), aber für jede Schulklasse unterschiedlich große Intercepts (random intercepts). Ihre Ausprägungen werden von den beiden Z-Variablen und den beiden W-Variablen beeinflusst.

Zur Interpretation des Gleichungssystems von Modell 27:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + \beta_{3jk}X1X2_{ijk} + e_{ijk} \quad (27.1)$$

Die Gleichung 27.1 enthält im Unterschied zu allen anderen, bislang erörterten Gleichungen der Modelle 20 bis 26 einen zusätzlichen Interaktionseffekt ( $\beta_{3jk}$ ), der von einer zusätzlichen Variablen (X1X2) ausgeht. In Abbildung 27.1 sind die zusätzlich ins Modell aufgenommenen Komponenten des Modells 27 innerhalb des rechteckigen Kastens zu finden. Die zusätzliche dichotome Interaktionsvariable wird als Produkt der beiden dichotomen Variablen X1 und X2 gebildet, sodass der geschätzte Effekt nur dann relevant wird, wenn sowohl X1 als auch X2 einen Wert von 1,00 aufweisen und somit auch die Interaktionsvariable einen Wert von 1,00 besitzt. Alle anderen Wertekombinationen (0/0, 0/1, 1/0) ergeben als Produkt immer einen Wert von 0,00.<sup>32</sup> Der geschätzte Interaktionseffekt ist als partieller Effekt zu verstehen, der unabhängig von den beiden Haupteffekten seine Wirkung ausübt.

Mplus: Within Level / MATH ON X1X2: 1,310 (p=0,671)

Mplus: Within Level / Residual Variances / MATH: 1417,997 (p=0,000)

$$\beta_{3jk} = \gamma_{30k} \quad (27.2d)$$

$$\gamma_{30k} = \delta_{300} \quad (27.3f)$$

Die beiden Gleichungen 27.2d und 27.3f definieren den Interaktionseffekt " $\beta_{3jk}$ " als fixierten Effekt, der über alle Analyseebenen hinweg den gleichen Wert aufweisen soll.

Ansonsten ist das formale Gleichungssystem von Modell 27 identisch mit dem Gleichungssystem von Modell 21 und Modell 25 (bis auf die oben erläuterten Modifikationen) und muss deshalb an dieser Stelle nicht noch einmal interpretiert werden.

---

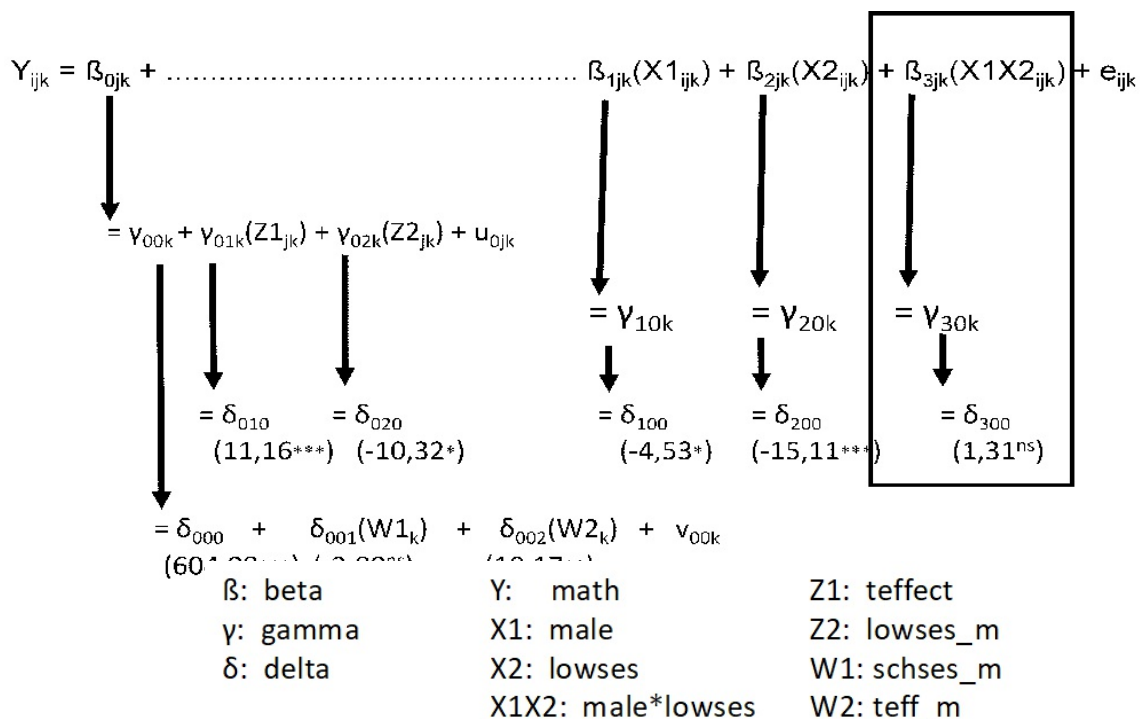
<sup>32</sup> Die multiplikative Konstruktion der Interaktionsvariablen gilt nur für Variablen, die beide auf der gleichen Analyseebene (hier: L1) liegen. Gehören die Variablen zu unterschiedlichen Analyseebenen (z.B. X2 und Z2), so muss ein "cross-level-effect" spezifiziert werden (vgl. dazu die Erläuterungen zu Modell 26).

Zur Interpretation der Schätzung von Modell 27:

(vgl. dazu auch die folgende Abbildung 27.3 sowie die folgenden Tabellen 27.1 und 27.2)

- 1) Der geschätzte Interaktionseffekt ist vom Betrag her sehr gering (1,31) und zudem auch noch hochgradig nicht-signifikant. Von daher ist nicht davon auszugehen, dass die Wertekombination “Geschlecht männlich” und “Herkunftsfamilie mit unterdurchschnittlichem SES” einen separaten sowie fixen Effekt auf die Höhe der MATH-Leistung ausübt.
- 2) Da das negative Testergebnis auf der Annahme beruht, dass es sich bei dem Interaktionseffekt um einen bereichübergreifenden, fixen Effekt handelt, sollte zur Absicherung des Ergebnisses ein zusätzliches Modell geschätzt werden, das die Annahme von einem fixen Interaktionseffekten aufgibt. Dies soll in der folgenden Modellschätzung 28 erfolgen.
- 3) Die Residual Varianz der Schätzung von Gleichung 27.1 ist sehr groß und höchst signifikant. Sie ist also mit extrem großer Wahrscheinlichkeit ungleich null. Wenn es die theoretischen bzw. analytischen Vorüberlegungen erlauben, können also noch weitere Prädiktoren zur Prognose der personalen Mathematikleistung herangezogen werden.
- 4) Ein Vergleich der AIC- und BIC-Werte von Modell 26 und Modell 27 zeigt, dass die Schätzung von Modell 27 eine bessere Anpassung aufweist als Modell 26 (vgl. Kap. B5).
- 5) Zu weiteren Erläuterungen der Schätzung von Modell 27 vgl. die Hinweise zu der Schätzung von Modell 21 (und dort vor allem Pkt. 7).

Abb. 27.3: Graphische Skizze des Gleichungssystems von Modell 27 mit Schätzergebnaten



Tab. 27.1: Ergebnisse der MLA-Schätzung von Modell 26: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L3-Intercepts ( $\delta_{000}$ ) MATH	604,089	0,000
L1-Slope ( $\delta_{100}$ ) MATH ON MALE	-4,535	0,016
L1-Slope ( $\delta_{200}$ ) MATH ON LOWSES	-15,11	0,000
L1-Slope ( $\delta_{300}$ ) MATH ON X1X2	1,310	0,671
L2-Slope ( $\delta_{010}$ ) MATH ON TEFFECT	11,160	0,000
L2-Slope ( $\delta_{020}$ ) MATH ON LOWSES_M	-10,329	0,013
L3-Slope ( $\delta_{001}$ ) MATH ON SCHSES_M	-3,805	0,314
L3-Slope ( $\delta_{002}$ ) MATH ON TEFF_M	10,172	0,001

Tab. 27.2: Ergebnisse der MLA-Schätzung von Modell 27: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH $\text{Var}(e_{ijk})$	1417,997	0.000
L2-Residualvarianz, MATH $\text{Var}(u_{0jk})$	92,167	0,003
L3-Residualvarianz, MATH $\text{Var}(v_{00k})$	46,997	0,032

Tab. 27.3: Mplus-Inputfile für Modell 27:

```

TITLE:
  Model 27 (three-level model / mit L1-Interaktionsvariable)
  mit random intercepts/fixed slopes,
  mit je zwei L1/2/3-Prädiktoren (dichotomisiert)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowsyes teffect
    lowsyes_m schses_m teff_m;

USEVARIABLES ARE
  teachid ! ist die L2-Clustervariable
  schcode ! ist die L3-Clustervariable
  math   ! ist die abh. Y-Variable
  lowsyes male ! sind L1-Prädiktoren
           ! male ist Rekodierung von female (s.u.)
  x1x2 ! ist die neue Interaktionsvariable (s.u.)
  z1 z2 w1 w2; ! werden alle über Mediansplit hergestellt (s.u.)

CLUSTER = schcode teachid;

MISSING ARE . ;

WITHIN = male lowsyes x1x2;
BETWEEN = (teachid) z1 z2 (schcode)w1 w2; ! in Klammern der Cluster-Identifizier

DEFINE:
  ! jetzt Dichotomisierung über Mediansplit:
  IF(teffect LE -.09)THEN z1 =0;
  IF(teffect GT -.09)THEN z1 =1;
  IF(lowsyes_m LE 0.33) THEN z2 =0;
  IF(lowsyes_m GT 0.33) THEN z2 =1;
  IF(schses_m LE 0.36)THEN w1 =0;
  IF(schses_m GT 0.36)THEN w1 =1;
  IF(teff_m LE 0.01) THEN w2 = 0;
  IF(teff_m GT 0.01) THEN w2 = 1;

  ! Definition von male
  IF(female == 1)THEN male=0;
  IF(female == 0)then male=1;

  ! Definition von Interaktionsvariable: x1x2
  x1x2=male*lowsyes;

ANALYSIS
  TYPE = THREELLEVEL; ! ACHTUNG: hier kein "random"

```

FORTSETZUNG —&gt;



Fortsetzung von Tabelle 27.3:

MODEL:

%WITHIN%

math ON male lowses x1 x2;    ! für L1  
                                  ! ohne female, denn female  
                                  ! wird nicht für das Modell gebraucht

%BETWEEN teachid%

math ON z1 z2;                ! für L2

%BETWEEN schcode%

math ON w1 w2;               ! für L3

OUTPUT:

SAMPSTAT;

Tab. 27.4: Mplus-Ausgabe von Modellschätzung 27 (ausgewählte Bereiche):

MODEL FIT INFORMATION				
Number of Free Parameters				11
Loglikelihood				
H0 Value				-14554.909
H0 Scaling Correction Factor for MLR				1.3333
H1 Value				-14554.909
H1 Scaling Correction Factor for MLR				1.3333
Information Criteria				
Akaike (AIC)				29131.818
Bayesian (BIC)				29197.393
Sample-Size Adjusted BIC				29162.443
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH	ON			
MALE	-4.535	1.875	-2.419	0.016
LOWSES	-15.110	2.315	-6.526	0.000
X1X2	1.310	3.086	0.424	0.671
Residual Variances				
MATH	1417.997	75.822	18.702	0.000
Between TEACHID Level				
MATH	ON			
Z1	11.160	2.519	4.430	0.000
Z2	-10.329	4.142	-2.494	0.013
Residual Variances				
MATH	92.167	31.212	2.953	0.003
Between SCHCODE Level				
MATH	ON			
W1	-3.805	3.777	-1.007	0.314
W2	10.172	2.974	3.420	0.001
Intercepts				
MATH	604.089	2.453	246.258	0.000
Residual Variances				
MATH	46.977	21.889	2.146	0.032

Modell 28:

Drei-Ebenen-Modell,

Random-Intercepts/Fixed+Random-Slopes,

mit je zwei L1/2/3-Prädiktoren (dichotom)

(“intercepts-as-outcomes model”)

(mit frei variierender L1-Interaktionsvariable)

L1-Gleichung:  $Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + \beta_{3jk}X1X2_{ijk} + e_{ijk}$  (28.1)

L2-Gleichungen:  $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}Z1_{jk} + \gamma_{02k}Z2_{jk} + u_{0jk}$  (28.2a)

$\beta_{1jk} = \gamma_{10k}$  (28.2b)

$\beta_{2jk} = \gamma_{20k}$  (28.2c)

$\beta_{3jk} = \gamma_{30k} + u_{3jk}$  (28.2d)

L3-Gleichungen:  $\gamma_{00k} = \delta_{000} + \delta_{001}W1_k + \delta_{002}W2_k + v_{00k}$  (28.3a)

$\gamma_{01k} = \delta_{010}$  (28.3b)

$\gamma_{02k} = \delta_{020}$  (28.3c)

$\gamma_{10k} = \delta_{100}$  (28.3d)

$\gamma_{20k} = \delta_{200}$  (28.3e)

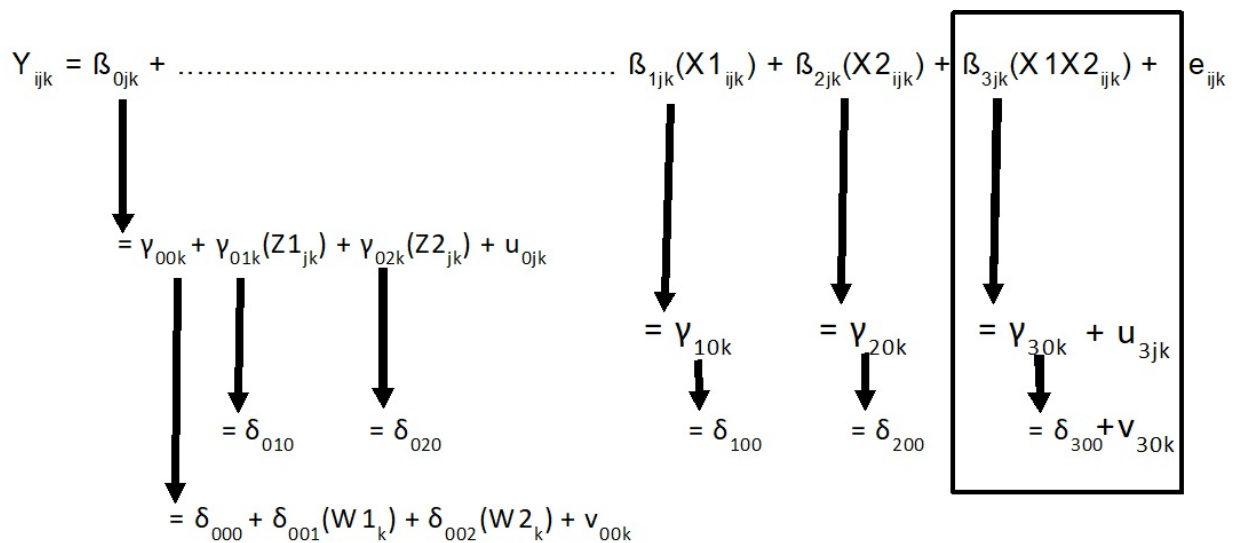
$\gamma_{30k} = \delta_{300} + v_{30k}$  (28.3f)

“fixed part”:  $\delta_{000}, \delta_{001}, \delta_{002}, \delta_{010}, \delta_{020}, \delta_{100}, \delta_{200}, \delta_{300}$

“random part”:  $\text{Var}(e_{ijk}), \text{Var}(u_{0jk}), \text{Var}(v_{00k}), \text{Var}(u_{3jk}), \text{Var}(v_{30k})$

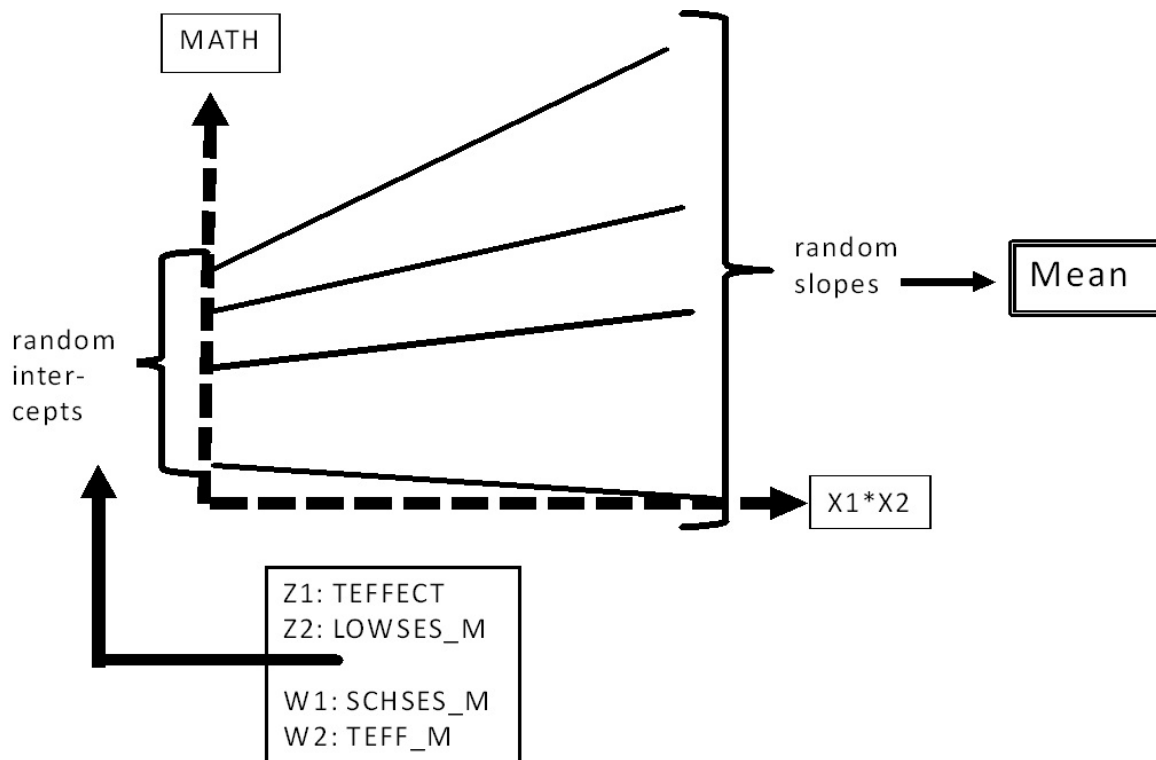
(Das Modell 28 entspricht dem Modell 27 mit der Modifikation, dass nunmehr der Interaktionseffekt von X1X2 (= MALE\*LOWSES) auf der L2- und L3-Ebene frei variiert (als “random effect”). Das Modell 28 orientiert sich an dem Modell 2 und dessen Formalisierung in Heck/Thomas 2015: 105-107, allerdings sind in Modell 28 alle Variablen dichotomisiert und nicht zentriert.)

Abb. 28.1: Graphische Skizze des Gleichungssystems von Modell 28



$\beta$ : beta	Y: math	Z1: teffect
$\gamma$ : gamma	X1: male	Z2: lowsyes_m
$\delta$ : delta	X2: lowsyes	W1: schses_m
	X1X2: male*lowsyes	W2: teff_m

Abb. 28.2: Stilisierte Darstellung der Modell-Logik von Modell 28

Zur Logik von Modell 28:

In der vorangegangenen Modellschätzung 27 wurde angenommen, dass der neu ins Modell aufgenommene Interaktionseffekt X1X2 nicht variiert, sondern klassen- bzw. schulübergreifend gültig ist. Demnach sollte der Interaktionseffekt für alle Schüler:innen (L1), in jeder Schulklasse (L2) und in jeder Schule (L3) gelten. Diese Annahme wird nunmehr in Modell 28 aufgegeben. Es soll untersucht werden, ob der Interaktionseffekt bedeutsam auf den Analyseebenen L2 und L3 variiert und somit als "random effect" zu verstehen ist. In Abbildung 28.1 befindet sich der neu spezifizierte Random-Interaktionseffekt in dem rechteckigen Kasten.

Die Abbildung 28.2 verdeutlicht die Logik von Modell 28. Das dargestellte Koordinatensystem (gestrichelte Linien) besteht aus der senkrechten Y-Achse (für die Variable MATH) und der waagrechten X-Achse (für die Variable X1\*X2). Der Zusammenhang zwischen Y- und X1X2-Variable wird für jede Schulklasse durch die Steigungslinie innerhalb des Koordinatensystems symbolisiert. Die Steigungslinien haben unterschiedliche Steigungsraten (random slopes). Zudem gibt es für jede Schulklasse unterschiedlich große Intercepts (random intercepts). Ihre Ausprägungen werden von den beiden Z-Variablen und den beiden W-Variablen beeinflusst.

Zur Interpretation des Gleichungssystems von Modell 28:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X1_{ijk} + \beta_{2jk}X2_{ijk} + \beta_{3jk}X1X2_{ijk} + e_{ijk} \quad (28.1)$$

Die Gleichung 28.1 enthält (wie schon die zuvor geschätzte Gleichung 27.1) die Interaktionsvariable X1X2. In den Erläuterungen zum Gleichungssystem 27 wurde bereits beschrieben, welche Annahmen mit der multiplikativen Konstruktion einer dichotomen Interaktionsvariablen verbunden sind. Auch in Modell 28 ist der geschätzte Interaktionseffekt als

partieller Effekt zu verstehen, der unabhängig von den beiden Haupteffekten seine Wirkung ausübt. Der Random Effect wird als Mittelwert aller schulspezifisch berechneten X1X2-Effekte ermittelt.

Mplus: Between SCHCODE level / Means SLOPE: -0,069 (p=0,982)

Mplus: Within Level / Residual Variances / MATH: 1417,573 (p=0,000)

$$\beta_{3jk} = \gamma_{30k} + u_{3jk} \quad (28.2d)$$

$$\gamma_{30k} = \delta_{300} + v_{30k} \quad (28.3f)$$

Die beiden Gleichungen 28.2d und 28.3f definieren den Interaktionseffekt “ $\beta_{3jk}$ ” als Random Effect, der über die L2 und L3 Analyseebenen hinweg variieren kann:

Mplus: Between TEACHID Level / Variances SLOPE: 0,984 (p=0,988)

Mplus: Between SCHCODE Level / Variances SLOPE: 0,970 (p=0,985)

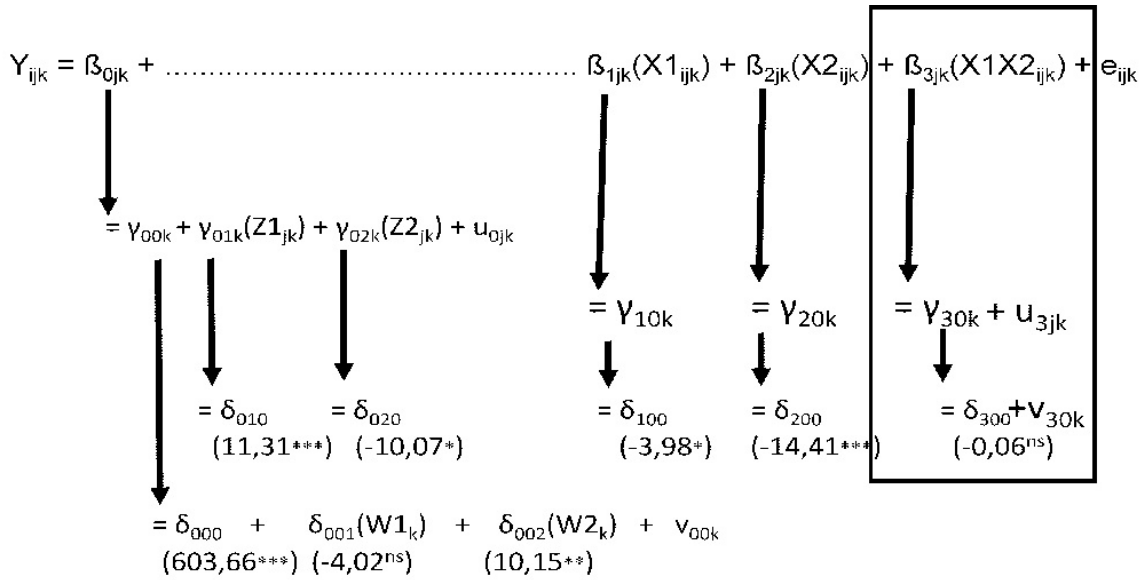
Ansonsten ist das formale Gleichungssystem von Modell 28 identisch mit dem Gleichungssystem von Modell 21 und Modell 25 (bis auf die oben erläuterten Modifikationen) und muss deshalb an dieser Stelle nicht noch einmal interpretiert werden.

#### Zur Interpretation der Schätzung von Modell 28:

(vgl. dazu auch die folgende Abbildung 28.3 sowie die folgenden Tabellen 28.1 und 28.2)

- 1) Auch in Modell 28 ist der geschätzte Interaktionseffekt vom Betrag her sehr gering (-0,06) und zudem auch noch hochgradig nicht-signifikant (p=0,970). Von daher ist nicht davon auszugehen, dass die Wertekombination “Geschlecht männlich” und “Herkunftsfamilie mit unterdurchschnittlichem SES” einen separaten Random-Effekt auf die Höhe der MATH-Leistung ausübt.
- 2) Da auch die Varianzen des nicht-signifikanten Interaktionseffekts auf L2 und L3 hochgradig nicht-signifikant sind (p=0,988 bzw. p=0,985) ist die Modellschätzung 28 (mit Random-Interaktionseffekt) der Modellschätzung 27 (mit Fixed-Interaktionseffekt) nicht deutlich überlegen. Beide Modellschätzungen liefern dementsprechend keine Hinweise darauf, den Interaktionseffekt in die Modellschätzung aufzunehmen. Allerdings sind die AIC/BIC-Werte von Modell 28 geringer als diejenigen von Modell 27. Trotzdem sollte in diesem Falle das Modell 27 bevorzugt werden (vgl. dazu Kap. B5).
- 3) Auch in der Modellschätzung 28 ist, wie in der Modellschätzung 27, die L1-Residual-Varianz der Schätzung (Gl. 27.1) sehr groß und höchst signifikant. Sie ist also mit extrem großer Wahrscheinlichkeit ungleich null. Es könnten also noch weitere Prädiktoren zur Prognose der personalen Mathematikleistung herangezogen werden, wenn dies die theoretischen bzw. analytischen Vorüberlegungen zur Modellspezifikation erlauben.
- 4) Zu weiteren Erläuterungen der Schätzung von Modell 28 vgl. die Hinweise zu der Schätzung von Modell 21.

Abb. 28.3: Graphische Skizze des Gleichungssystems von Modell 28 mit Schätzergebnaten



$\beta$ : beta                      Y: math                      Z1: teffect  
 $\gamma$ : gamma                    X1: male                    Z2: lowsese\_m  
 $\delta$ : delta                      X2: lowsese                W1: schses\_m  
    X1X2: male\*lowese        W2: teff\_m

Tab. 28.1: Ergebnisse der MLA-Schätzung von Modell 28: “fixed coefficients”

Variablen	Schätzung	Sign.Niveau (p) zweiseitig
L3-Intercepts ( $\delta_{000}$ ) MATH	603,666	0,000
L1-Slope ( $\delta_{100}$ ) MATH ON MALE	-3,985	0,035
L1-Slope ( $\delta_{200}$ ) MATH ON LOWSES	-14,413	0,000
L3-Slope, Mean ( $\delta_{300}$ ) MATH ON X1X2	-0,069	0,982
L2-Slope ( $\delta_{010}$ ) MATH ON TEFFECT	11,312	0,000
L2-Slope ( $\delta_{020}$ ) MATH ON LOWSES_M	-10,071	0,015
L3-Slope ( $\delta_{001}$ ) MATH ON SCHSES_M	-4,024	0,289
L3-Slope ( $\delta_{002}$ ) MATH ON TEFF_M	10,157	0,001

Tab. 28.2: Ergebnisse der MLA-Schätzung von Modell 28: “random coefficients”

Varianzen	Schätzung	Sign.Niveau (p) zweiseitig
L1-Residualvarianz, MATH $\text{Var}(e_{ijk})$	1417,573	0.000
L2-Residualvarianz, MATH $\text{Var}(u_{ojk})$	92,974	0,004
L3-Residualvarianz, MATH $\text{Var}(v_{00k})$	47,056	0,033
L2-Slopevarianz, X1X2 $\text{Var}(u_{3jk})$	0,984	0,988
L3-Slopevarianz, X1X2 $\text{Var}(v_{30k})$	0,970	0,985

Tab. 28.3: Mplus-Inputfile für Modell 28

```

TITLE:
  Model 28 (three-level model)
  mit random intercepts/fixed+random slopes (random slopes für beta X1X2)
  mit je zwei L1/2/3-Prädiktoren (dichotomisiert)

DATA:
  FILE IS e:\data\ch4three.dat;
  FORMAT IS 6f8.0,4f8.2;

VARIABLE:
  NAMES ARE
    person teachid schcode math female lowses teffect
    lowses_m schses_m teff_m;
  USEVARIABLES ARE
    teachid      ! ist die L2-Clustervariable
    schcode      ! ist die L3-Clustervariable
    math         ! ist die abh. Y-Variable
    ! ohne female, denn female wird
    ! für das Modell nicht gebraucht
    lowses male ! sind L1-Prädiktoren
    z1 z2 w1 w2 ! werden alle über Mediansplit hergestellt
    x1x2;       ! ist die neue Interaktionsvariable (s.u.)

  CLUSTER = schcode teachid;

  MISSING ARE . ;

  WITHIN = male lowses x1x2;

  BETWEEN = (teachid) z1 z2 (schcode)w1 w2;
             ! in Klammern ist der jeweilige Cluster-Identifizier

DEFINE:
  ! jetzt Dichotomisierung über Mediansplit:
  IF(teffect LE -.09)THEN z1 =0;
  IF(teffect GT -.09)THEN z1 =1;
  IF(lowses_m LE 0.33) THEN z2 =0;
  IF(lowses_m GT 0.33) THEN z2 =1;
  IF(schses_m LE 0.36)THEN w1 =0;
  IF(schses_m GT 0.36)THEN w1 =1;
  IF(teff_m LE 0.01) THEN w2 = 0;
  IF(teff_m GT 0.01) THEN w2 = 1;

  if(female == 1)then male=0;
  if(female == 0)then male=1;

  ! jetzt mit Interaktionsvariable: x1x2
  x1x2=male*lowses;

ANALYSIS:
  TYPE = THREELEVEL random;

```

FORTSETZUNG →



Fortsetzung von Tabelle 28.3:

MODEL:

```
%WITHIN%  
  math ON male lowses;  
  slope | math on x1x2;  ! für L1  
  
%BETWEEN teachid%  
  math ON z1 z2;      ! für L2  
  
%BETWEEN schcode%  
  math ON w1 w2;      ! für L3
```

OUTPUT:

```
SAMPSTAT;
```

Tab. 28.4: Mplus-Ausgabe von Modellschätzung 28 (ausgewählte Bereiche)

MODEL FIT INFORMATION				
Number of Free Parameters				13
Loglikelihood				
H0 Value				-14555.037
H0 Scaling Correction Factor for MLR				1.2811
Information Criteria				
Akaike (AIC)				29136.074
Bayesian (BIC)				29213.571
Sample-Size Adjusted BIC				29172.266
(n* = (n + 2) / 24)				
MODEL RESULTS				
			Two-Tailed	
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
MATH	ON			
MALE	-3.985	1.885	-2.114	0.035
LOWSES	-14.413	2.321	-6.209	0.000
Residual Variances				
MATH	1417.573	77.523	18.286	0.000
Between TEACHID Level				
MATH	ON			
Z1	11.312	2.515	4.498	0.000
Z2	-10.071	4.144	-2.430	0.015
Variances				
SLOPE	0.984	63.557	0.015	0.988
Residual Variances				
MATH	92.974	32.349	2.874	0.004
Between SCHCODE Level				
MATH	ON			
W1	-4.024	3.793	-1.061	0.289
W2	10.157	3.028	3.354	0.001
Means				
SLOPE	-0.069	3.142	-0.022	0.982
Intercepts				
MATH	603.666	2.479	243.479	0.000
Variances				
SLOPE	0.970	51.025	0.019	0.985
Residual Variances				
MATH	47.056	22.058	2.133	0.033

**F Tabellarischer Modellvergleich**

random intercepts	fixed slopes	random slopes	X-Var.	Z-Var.	W-Var.	dicho. Var.	cross level effects	intercepts als outcome	slopes als outcome	Modell
<b>Zwei-Ebenen-Modelle:</b>										
X										M1
X	X		1							M2
X	X		2							M3
X				1		1		X		M4
X				1				X		M5
X		X	1							M6
X		X	2			1				M7
X		X	2			2				M8
X	X	X	2			2				M9
X		X	1	1		1		X		M10
X		X	1	1		1	X	X	X	M11
X		X	2	1		3	X	X	X	M12
X		X	2	1		3	X	X	X	M13
X	X	X	2	1		3	X	X	X	M14
<b>Drei-Ebenen-Modelle:</b>										
X										M20
X	X		2	2	2	2		X		M21
X	X	X	2	2	2	2		X		M22
X	X	X	2	2	2	2	X	X	X	M23
X	X	X	2	2	2	6	X	X	X	M24
X	X	X	2	2	2	6		X		M25
X	X	X	2	2	2	2	X	X	X	M26
X	X		3	2	2	3		X		M27
X	X	X	3	2	2	3		X		M28

**G Gebräuchliche Notation von Komponenten der Mehrebenenanalyse**

- L1 Level 1: unterste Analyseebene mit nicht-kontextbezogenen und nicht-aggregierten L1-Variablen (wie z.B. die Variablen "Alter" oder "Geschlecht" bei einzelnen Befragten)
- L2 Level 2: Makro/Kontext/Cluster-Ebene mit L2-Variablen (wie z.B. das Bruttoinlandsprodukt (BIP bzw. GDP) eines Landes/Staates)
- L3 Level 3: Makro/Kontext/Cluster-Ebene mit L3-Variablen (wie z.B. neben den Merkmalen von Schüler:innen (L1) und den Merkmalen von Schulklassen (L2) die Merkmale von Schulen: deren Größe, deren Alter usw.)
- $Y_{ijk}$  abh. L1-Variable "Y" (z.B. "die schulische Leistung") mit einem Messwert für jede Person "i" in jeder Schulklasse "j" und in jeder Schule "k"
- $X_{ijk}$  unabh. L1-Prädiktor/L1-Variable "X" (z.B. "Alter") mit einem Messwert für jede Person "i" in jeder Schulklasse "j" und in jeder Schule "k"
- $Z_{jk}$  unabh. L2-Prädiktor/L2-Variable "Z" (z.B. "Größe der Schulklasse") mit einem Messwert für jede Schulklasse "j" in jeder Schule "k"
- $W_k$  unabh. L3-Prädiktor/L3-Variable "W" (z.B. "Größe der Schule") mit einem Messwert für jede Schule "k"
- $\beta_{0jk}$  Intercept/Y-Achsenabschnitt in einer L1-Regressionsschätzung von Y mit X-Werten
- $\beta_{1jk}$  Slope/Steigungskoeffizient/Regressionskoeffizient in einer L1-Regressionsschätzung von Y mit X1-Werten
- $e_{ijk}$  Residualwert in einer L1-Regressionsschätzung für eine Person "i" in den Clustern "j" und "k" (= Differenz zwischen geschätztem und tatsächlichen Wert von Y)
- $\text{Var}(e_{ijk})$  Varianz der L1-Residualwerte "e<sub>ijk</sub>" (s.o.) (= im Modell unerklärt verbliebene L1-Varianz)
- $\gamma_{00k}$  Intercept/Y-Achsenabschnitt in einer L2-Regressionsschätzung des L1-Random Intercept  $\beta_{0jk}$  mit Z-Variablen
- $\gamma_{01k}$  Slope/Steigungskoeffizient/Regressionskoeffizient in einer L2-Regressionsschätzung des L1-Random Intercept  $\beta_{0jk}$  mit Z1-Messwerten
- $\gamma_{10k}$  Intercept/Achsenabschnitt in einer L2-Regressionsschätzung des L1-Random Slope  $\beta_{1jk}$  mit Z-Variablen
- $\gamma_{15k}$  Slope/Steigungskoeffizient/Regressionskoeffizient in einer L2-Regressionsschätzung des L1-Random Slope  $\beta_{1jk}$  mit Z1-Messwerten
- $u_{0jk}$  Residualwert in einer L2-Regressionsschätzung des L1-Random Intercept  $\beta_{0jk}$  mit Z-Variablen

$u_{1jk}$  Residualwert in einer L2-Regressionsschätzung des L1-Random Slopes  $\beta_{1jk}$  mit Z-Variablen

$\text{Var}(u_{0jk})$  Varianz der Residualwerte (s.o.), Varianz des

$\text{Var}(u_{1jk})$  Varianz der Residualwerte (s.o.)

$\delta_{000}$  Intercept/Achsenabschnitt in einer L3-Regressionsschätzung des L2-Random Intercept  $\gamma_{00k}$  mit W-Variablen

$\delta_{001}$  Slope/Steigungskoeffizient/Regressionskoeffizient in einer L3-Regressionsschätzung des L2-Random Intercept  $\gamma_{00k}$  mit W1-Messwerten

$\delta_{010}$  Intercept/Achsenabschnitt in einer L3-Regressionsschätzung des L2-Random Slope  $\gamma_{01k}$  (mit W-Variablen)

$\delta_{011}$  Slope/Steigungskoeffizient/Regressionskoeffizient in einer L3-Regressionsschätzung des L2-Random Slopes  $\gamma_{00k}$  mit W1-Messwerten

$v_{00k}$  Residualwert in einer L3-Regressionsschätzung des L2-Random Intercepts  $\gamma_{00k}$

$v_{01k}$  Residualwert in einer L3-Regressionsschätzung des L2-Random Slopes  $\gamma_{01k}$

$\text{Var}(v_{00k})$  Varianz der Residualwerte (s.o.)

$\text{Var}(v_{01k})$  Varianz der Residualwerte (s.o.)

## H Literatur

- Bickel, R., 2007: *Multilevel Analysis for Applied Research. It's Just Regression*. New York (Guilford).
- Eid, M. et al., 2013: *Statistik und Forschungsmethoden*. Weinheim (Beltz).
- Field, A. et.al., 2012: *Discovering Statistics Using R*. London (Sage).
- Finch, W.H./Bolin, J.E., 2017: *Multilevel Modeling Using Mplus*. Boca Raton (Chapman & Hall/CRC).
- Geiser, C., 2010: *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden (VS Verlag).
- Hadler, M., 2004: Die Mehrebenen-Analyse. Ihre praktische Anwendung und theoretische Annahmen. In: *Österreichische Zeitschrift für Soziologie*, Vol. 29, Nr. 1, S. 53-74.
- Heck, R.H./Thomas, S.L., 2015: *An Introduction to Multilevel Modeling Techniques. MLM und SEM Approaches Using Mplus*. 3rd ed., New York (Routledge).
- Hox, J.J./Moerbeek, M./van de Schoot, R., 2018: *Multilevel Analysis. Techniques and Applications*. 3rd ed., New York (Routledge).
- Huang, F.L., 2018: Multilevel Modeling Myths. *School Psychology Quarterly*, p. 1-18.
- Luhmann, M., 2015: *R für Einsteiger. Einführung in die Statistiksoftware für die Sozialwissenschaften*, 4. Aufl., Weinheim/Basel (Beltz).
- Maas, C. J./Hox, J. J., 2005: Sufficient sample sizes for multilevel modeling. *Methodology* 1: 86–92.
- Moineddin, R./Matheson, F. I./Glazier, R. H., 2007: A simulation study of sample size for multilevel logistic regression models. *BMC Medical Research Methodology* 7.
- Muthén, L. K./Muthén, B. O., 2002: How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling* 9: 599–620.
- Paccagnella, O., 2011: Sample size and accuracy of estimates in multilevel models: New simulation results. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 7: 111–120.
- Poetschke, M., 2006: Mehrebenenanalyse, S. 167-179 in: Behnke, J. et al. (eds.), *Methoden der Politikwissenschaft*. Baden-Baden (Nomos).
- Poetschke, M., 2014: Aktuelle Probleme der Modellierung von Mehrebenen-Daten. S. 219-240 in: Friedrichs, J./Nonnemacher, A. (eds.), *Soziale Kontexte und soziale Mechanismen*, Wiesbaden (Springer VS)
- Schoeneberger, J. A., 2016: The impact of sample size and other factors when estimating multilevel logistic models. *The Journal of Experimental Education* 84: 373–397.
- Urban, D./Mayerl, J., 2014: *Strukturgleichungsmodellierung. Ein Ratgeber für die Praxis*. Wiesbaden (Springer VS).
- Urban, D./Mayerl, J., 2018: *Angewandte Regressionsanalyse: Theorie, Technik und Praxis*. 5. Aufl., Wiesbaden (Springer VS).

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 51/2022 Mehrebenenanalyse: Die wichtigsten 23 Modelle zur Analyse von dichotomen und metrischen Zwei- und Drei-Ebenen-Effekten (unter Verwendung von Mplus).  
Dieter Urban
- No. 50/2020 Verfahren der Multiplen Imputation bei Schätzung von Strukturgleichungsmodellen mit latenten Variablen. Ein systematischer Vergleich mittels Monte-Carlo-Simulationen.  
Andreas Wahl, Dieter Urban
- No. 49/2020 Psychische Gesundheit: Eine heikle Thematik in der empirischen Umfrageforschung?  
Zur Validierung des Crosswise Modells.  
Anna Götze, Andreas Wahl
- No. 48/2020 Analyse und Schätzung von Mehrgruppen-Strukturgleichungsmodellen mittels SPSS und EQS. Eine praxisnahe Anleitung.  
Andreas Wahl, Dieter Urban
- No. 47/2020 Machine Learning basierte Response Style Identifikation:  
eine simulations-statistische Pilotstudie.  
Thomas Krause
- No. 46/2020 Informationskriterien zur latenten Klassenidentifikation für Multiple Indicator Growth Mixture Models.  
Thomas Krause
- No. 45/2017 Die Angst vor dem Fremden. Ein Vignettenexperiment zur Bedrohungswahrnehmung.  
Ergebnisse eines studentischen Projektseminars.  
Alexander Geisler, Alexandra Möck, Rebecca Nell, Svenja Wiechmann, Thomas Krause
- No. 44/2016 Regressionsanalyse bei fehlenden Variablenwerten (missing values): Imputation oder Nicht-Imputation? Eine Anleitung für die Regressionspraxis mit SPSS.  
Dieter Urban, Jochen Mayerl, Andreas Wahl
- No. 43/2014 Meta-Analyse mit Mehrebenenmodellen unter Verwendung von HLM.  
Joachim Fiebig, Dieter Urban
- No. 42/2013 Panelanalyse mit Mehrebenenmodellen. Eine anwendungsorientierte Einführung.  
Thomas Krause, Dieter Urban
- No. 41/2012 Meta-Analyse und "publication bias". Eine empirisch-statistische Fallstudie.  
Joachim Fiebig, Gerrit Rüdebusch, Dieter Urban
- No. 40/2011 Pädosexuelle Viktimisierung und pädosexuelle Straffälligkeit. Ein „harter“ empirischer Hypothesentest.  
Dieter Urban, Joachim Fiebig
- No. 39/2010 Binär-logistische Regressionsanalyse.  
Grundlagen und Anwendung für Sozialwissenschaftler.  
Jochen Mayerl, Dieter Urban
- No. 38/2010 Das Working-Poor-Problem in Deutschland.  
Empirische Analysen zu den Ursachen von Armut trotz Erwerbstätigkeit.  
Leonie Hellmuth, Dieter Urban
- No. 37/2010 Der Bystander-Effekt in alltäglichen Hilfsituationen: Ein nicht-reaktives Feldexperiment.  
Katrin Alle, Jochen Mayerl

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 36/2008    Berufseintritt und Berufssituation von Soziologieabsolventen der Universität Stuttgart. Deskriptive Ergebnisse einer Absolventenbefragung aus dem Jahr 2007.  
Jochen Mayerl, Dieter Urban
- No. 35/2007    Antwortlatenzzeiten in TRA-Modellen. Zur statistischen Erklärung von (Geld)-Spendenverhalten.  
Dieter Urban, Jochen Mayerl
- No. 34/2006    Zur soziologischen Erklärung individuellen Geldspendens. Eine Anwendung und Erweiterung der Theory of Reasoned Action unter Verwendung von Antwortlatenzzeiten in einem Mediator-Moderator-Design.  
Jochen Mayerl
- No. 33/2006    Ist Bildung gleich Bildung? Der Einfluss von Schulbildung auf ausländerablehnende Einstellungen in verschiedenen Alterskohorten.  
Klaus Hadwiger
- No. 32/2005    Analyzing cognitive processes in CATI-Surveys with response latencies: An empirical evaluation of the consequences of using different baseline speed measures.  
Jochen Mayerl, Piet Sellke, Dieter Urban
- No. 31/2005    Persönliche Netzwerke in der Systemtheorie.  
Jan A. Fuhse
- No. 30/2004    Erhöht ein Opfer-Täter-Zyklus das Risiko, Sexualstraftaten als pädosexuelle Straftaten zu begehen? Ergebnisse einer ereignisanalytischen Pilotstudie.  
Dieter Urban, Heiko Lindhorst
- No. 29/2003    Können Nonattitudes durch die Messung von Antwortreaktionszeiten ermittelt werden? Eine empirische Analyse computergestützter Telefoninterviews.  
Jochen Mayerl
- No. 28/2003    Systeme, Netzwerke, Identitäten. Die Konstitution sozialer Grenzziehungen am Beispiel amerikanischer Straßengangs.  
Jan A. Fuhse
- No. 27/2002    Probleme bei der Messung individueller Veränderungsraten. 13 empirisch und methodisch induzierte Effekte, die es schwierig machen, Veränderungen von generalisierten Bewertungen zu ermitteln.  
Dieter Urban
- No. 26/2002    Vertrauen und Risikoakzeptanz. Zur Relevanz von Vertrauen bei der Bewertung neuer Technologien.  
Martin Slaby, Dieter Urban
- No. 25/2002    Risikoakzeptanz als individuelle Entscheidung. Zur Integration der Risikoanalyse in die nutzentheoretische Entscheidungs- und Einstellungsforschung.  
Martin Slaby, Dieter Urban
- No. 24/2001    Ist Glück ein affektiver Sozialindikator subjektiven Wohlbefindens? Dimensionen des subjektiven Wohlbefindens und die Differenz zwischen Glück und Zufriedenheit.  
Jochen Mayerl



## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 23/2001    Religiosität und Profession. Longitudinale Analysen zur Entwicklung des religiösen Selbstbildes bei Erzieherinnen.  
Heiko Lindhorst
- No. 22/2001    Differentielle Technikakzeptanz, oder: Nicht immer führt die Ablehnung einer Technik auch zur Ablehnung ihrer Anwendungen. Eine nutzentheoretische und modell-statistische Analyse.  
Martin Slaby, Dieter Urban
- No. 21/2001    Unser „wir“ - ein systemtheoretisches Modell von Gruppenidentitäten.  
Jan A. Fuhse
- No. 20/2000    Längsschnittanalysen mit latenten Wachstumskurvenmodellen in der politischen Sozialisationsforschung.  
Dieter Urban
- No. 19/2000    Die Zufriedenheit von Stuttgarter Studierenden mit ihrer Lebens- und Wohnsituation. Erste deskriptive Ergebnisse einer sozialwissenschaftlichen Studie zu allgemeinen und bereichsspezifischen Zufriedenheiten der Studierenden des Campus Vaihingen und des Campus Hohenheim.  
Projektgruppe Campus: Slaby, M.; Grund, R.; Mayerl, J.; Noak, T.; Payk, B.; Sellke, P.; Urban, D.; Zudrell, I.
- No. 18/1999    Role Models and Trust in Socio-Political Institutions: A Case Study in Eastern Germany, 1992-96.  
Joachim Singelmann, Toby A. Ten Ayck, Dieter Urban
- No. 17/1998    Zur Interaktion zwischen Befragten und Erhebungsinstrument. Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf.  
Martin Slaby
- No. 16/1998    Technikeinstellungen: gibt es die überhaupt? Ergebnisse einer Längsschnittanalyse von Bewertungen der Gentechnik.  
Dieter Urban
- No. 15/1998    Bewertende Einstellungen zur Gentechnik: ihre Form, ihre Inhalte und ihre Dynamik. Kurzbericht zu Ergebnissen des Forschungsprojektes „Einstellungen zur Gentechnik“.  
Dieter Urban, Uwe Pfenning, Joachim Allhoff
- No. 14/1997    Politische Sozialisation im Transformationsprozeß: Die Entwicklung demokratiebezogener Einstellungen von ostdeutschen Jugendlichen und deren Eltern zwischen 1992 und 1996.  
Barbara Schmidt, Dieter Urban, Joachim Singelmann
- No. 13/1997    Die soziale Vererbung von „Ausländerfeindlichkeit“. Eine empirische Längsschnittanalyse der intra- und intergenerativen Transmission von sozialen Einstellungen.  
Dieter Urban und Joachim Singelmann
- No. 12/1997    Gentechnik: „Fluch oder Segen“ versus „Fluch und Segen“. Bilanzierende und differenzierende Bewertungen der Gentechnik in der öffentlichen Meinung.  
Dieter Urban und Uwe Pfenning

## **SISS:**

### **Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**

<https://www.sowi.uni-stuttgart.de/forschung/schriftenreihe-des-instituts-fuer-sozialwissenschaften-siss/>

- No. 11/1996 Einstellungsmessung oder Einstellungsgenerierung? Die Bedeutung der informationellen Basis bei Befragten für die empirische Rekonstruktion von Einstellungen zu gentechnischen Anwendungen.  
Martin Slaby
- No. 10/1996 Jugend und Politik im Transformationsprozeß. Eine Fallstudie zur Stabilität und Veränderung von politischen Einstellungen bei ostdeutschen Jugendlichen zwischen 1992 und 1995.  
Dieter Urban, Joachim Singelmann, Helmut Schröder
- No. 09/1995 Social Indicators in a Nonmetropolitan County: Testing the Representativeness of a Regional Nonrandom Survey in Eastern Germany.  
Dieter Urban, Joachim Singelmann
- No. 08/1995 Handbuch zur empirischen Erhebung von Einstellungen/Kognitionen zur Bio- und Gentechnologie (inklusive Diskette)  
(zweite, überarbeitete und erweiterte Auflage)  
Uwe Pfenning, Dieter Urban, Volker Weiss
- No. 07/1995 Latent Attitude Structures Directing the Perception of New Technologies.  
An Application of SEM-Methodology to the Construction of Attitude Measurement Models Related to Technologies of Prenatal Genetic Engineering and Testing.  
Dieter Urban
- No. 06/1995 Regionale Strukturprobleme.  
Sammelband der Beiträge zum Symposium vom 13. und 14. Oktober 1994.  
Frank C. Englmann (Hrsg.)
- No. 05/1995 Die Bedeutung Neuronaler Netze in der Ökonomie.  
Hermann Schnabl
- No. 04/1994 Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht aus den neuen Bundesländern.  
Udo Kornblum
- No. 03/1994 Neue württembergische Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht.  
Udo Kornblum
- No. 02/1994 Report on the German Machine Tool Industry.  
Frank C. Englmann, Christian Heyd, Daniel Köstler, Peter Paustian  
with the assistance of Susanne Baur and Peter Bergmann
- No. 01/1994 "Vertrauen" - soziologisch betrachtet. Ein Beitrag zur Analyse binärer Interaktionssysteme.  
Peter Antfang, Dieter Urban