

Lineare Strukturgleichungsmodelle

- 1 Korrelation und Kausalität
- 2 Grundsätzliches
- 3 Pfaddiagramme und lineare Strukturgleichungen
- 4 Struktur- und Messmodell
- 5 Modellspezifikation
- 6 Parameterschätzungen
- 7 Beurteilung der Schätzergebnisse
- 8 LSM mit latenten Variablen
- 9 Software – AMOS

Literatur: Rudolf & Müller, S. 337-390

Korrelation und Kausalität

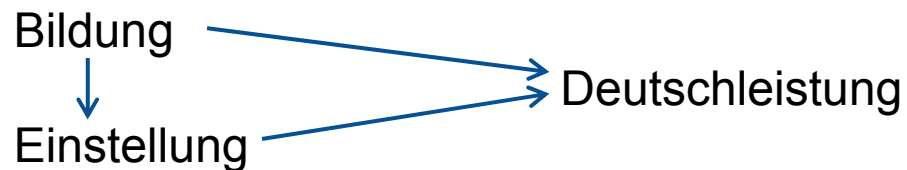
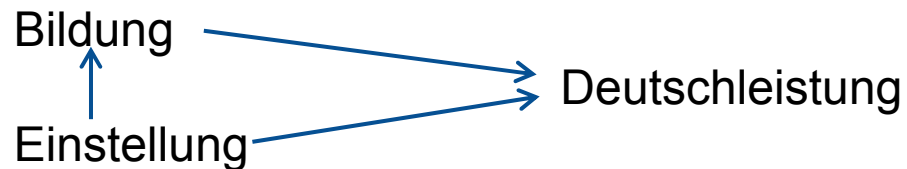
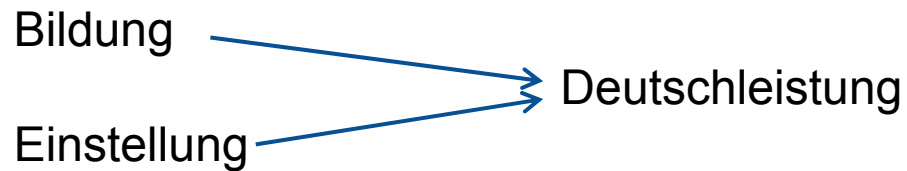
Korrelationen

| | | Bildungsniveau der Eltern | Einstellung der Eltern zur Schule | Ergebnis Deutsch 4. Klasse | Ergebnis Deutsch 10. Klasse |
|-----------------------------------|--------------------------|---------------------------|-----------------------------------|----------------------------|-----------------------------|
| Bildungsniveau der Eltern | Korrelation nach Pearson | 1 | ,529** | ,788** | ,691** |
| | Signifikanz (2-seitig) | | ,000 | ,000 | ,000 |
| | N | 120 | 120 | 120 | 120 |
| Einstellung der Eltern zur Schule | Korrelation nach Pearson | ,529** | 1 | ,764** | ,688** |
| | Signifikanz (2-seitig) | ,000 | | ,000 | ,000 |
| | N | 120 | 120 | 120 | 120 |
| Ergebnis Deutsch 4. Klasse | Korrelation nach Pearson | ,788** | ,764** | 1 | ,904** |
| | Signifikanz (2-seitig) | ,000 | ,000 | | ,000 |
| | N | 120 | 120 | 120 | 120 |
| Ergebnis Deutsch 10. Klasse | Korrelation nach Pearson | ,691** | ,688** | ,904** | 1 |
| | Signifikanz (2-seitig) | ,000 | ,000 | ,000 | |
| | N | 120 | 120 | 120 | 120 |

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

➤ Korrelation und Kausalität

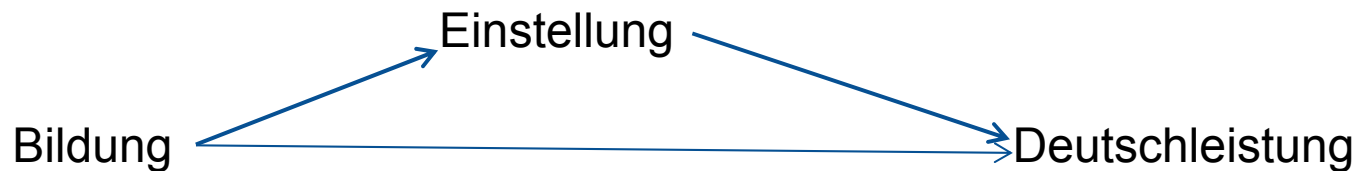
Auswahl an alternativen Erklärungsmöglichkeiten für die Korrelationen zwischen Bildung, Einstellung und Deutsch 4. Klasse:



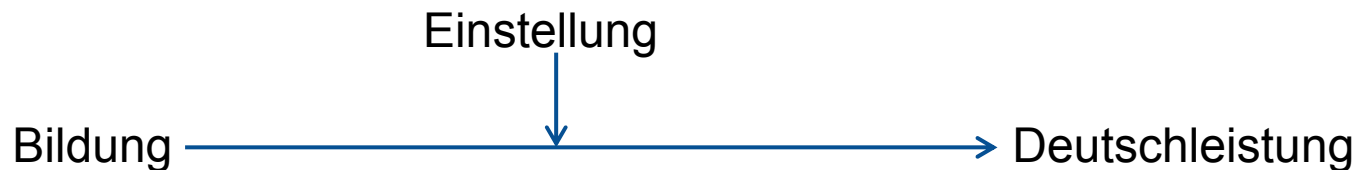
➤ Korrelation und Kausalität

Auswahl an alternativen Erklärungsmöglichkeiten für die Korrelationen zwischen Bildung, Einstellung und Deutsch 4. Klasse:

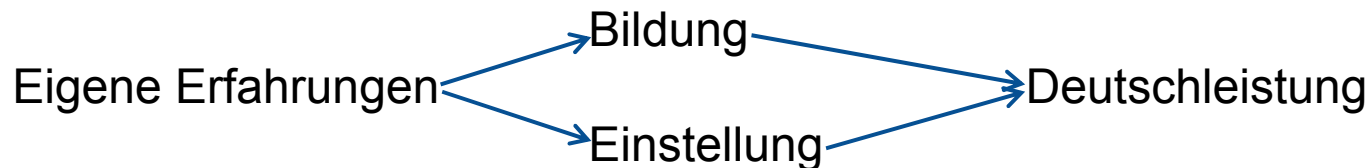
Mediation:








Moderation:



Drittvariablen:



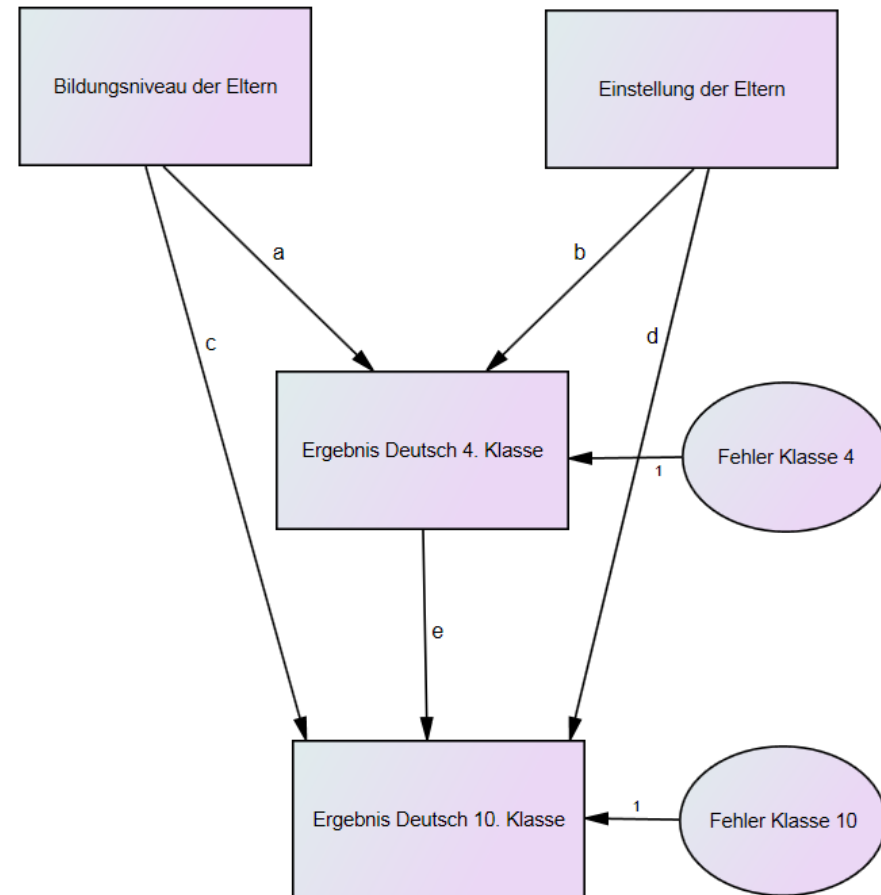
Korrelation und Kausalität

-  sehr unterschiedliche Modelle zur Erklärung der Korrelationen möglich
-  Bewertung unterschiedlicher Modelle in einfachen Fällen über partielle Korrelationen, semipartielle Korrelationen u.ä. möglich
-  Bewertung unterschiedlicher Modelle in komplexeren Situationen: Pfadmodelle bzw. lineare Strukturgleichungsmodelle
-  i.d.R. keine eindeutigen Ergebnisse aus der Datenanalyse möglich
-  Inhaltliche Modellbildung von entscheidender Bedeutung

➤ Pfadmodelle: Grundideen

Grundlage: Inhaltliche Hypothesen

Umsetzung in Grafiken...



... und in Gleichungen
(Variablen verkürzt bezeichnet)

$$\text{Deutsch 4} = a \cdot \text{Bildung} + b \cdot \text{Einstellung} + \text{Fehler 4}$$

$$\text{Deutsch 10} = c \cdot \text{Bildung} + d \cdot \text{Einstellung} + e \cdot \text{Deutsch 4} + \text{Fehler 10}$$

Pfadmodelle: Grundideen

Einige wichtige zusätzliche Möglichkeiten von Pfadmodellen gegenüber multipler linearer Regressionsanalyse:

- Modellfehler werden explizit modelliert
- Kovarianzen (Korrelationen) zwischen Prädiktoren sind zugelassen und können modelliert werden (siehe folgende Folie)
- „Pfade“ können abgebildet werden, d.h. Variablen können sowohl als Prädiktoren als auch als Kriterien im Modell wirken
- Erweiterung auf Strukturgleichungsmodelle mit latenten Variablen möglich

Pfadmodelle: Pfadkoeffizienten

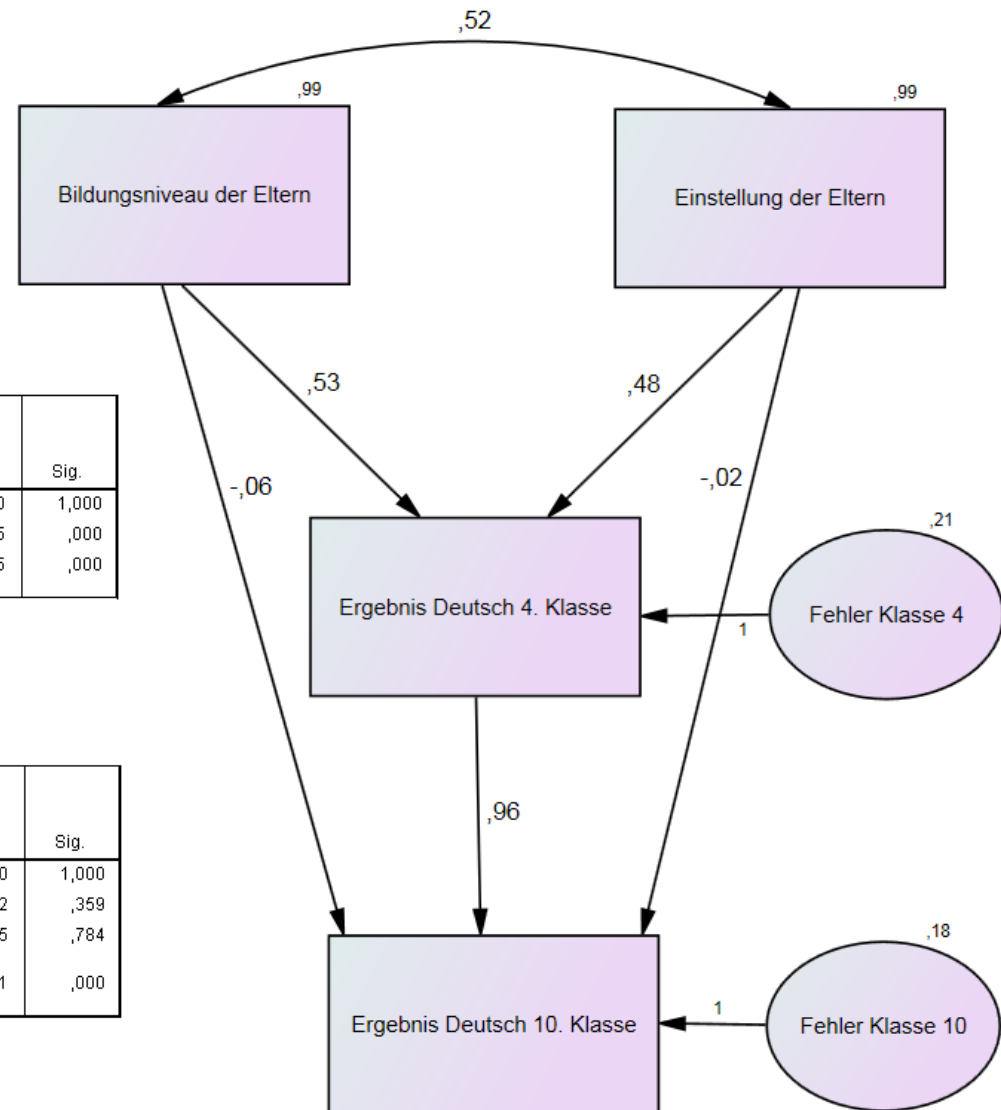
Berechnung der Pfadkoeffizienten im Ergebnis multipler Regressionsanalysen (Variablen liegen z-standardisiert vor):

Multiple Regression 1: AV: Deutsch 4. Klasse

| Modell | | Nicht standardisierte Koeffizienten | | Standardisierte Koeffizienten | T | Sig. |
|--------|-----------------------------------|-------------------------------------|----------------|-------------------------------|--------|-------|
| | | Regressionskoeffizient B | Standardfehler | Beta | | |
| 1 | (Konstante) | 7,828E-17 | ,042 | | ,000 | 1,000 |
| | Bildungsniveau der Eltern | ,533 | ,050 | ,533 | 10,625 | ,000 |
| | Einstellung der Eltern zur Schule | ,482 | ,050 | ,482 | 9,625 | ,000 |

Multiple Regression 2: AV: Deutsch 10. Klasse

| Modell | | Nicht standardisierte Koeffizienten | | Standardisierte Koeffizienten | T | Sig. |
|--------|-----------------------------------|-------------------------------------|----------------|-------------------------------|--------|-------|
| | | Regressionskoeffizient B | Standardfehler | Beta | | |
| 1 | (Konstante) | -5,905E-17 | ,039 | | ,000 | 1,000 |
| | Bildungsniveau der Eltern | -,060 | ,065 | -,060 | -,922 | ,359 |
| | Einstellung der Eltern zur Schule | -,017 | ,062 | -,017 | -,275 | ,784 |
| | Ergebnis Deutsch 4. Klasse | ,965 | ,086 | ,965 | 11,251 | ,000 |



➤ Pfadmodelle: Typen von Effekten

Direkte kausale Effekte

Zum Beispiel direkter kausaler Effekt von **Bildung** auf **Deutsch 10**: -0.06

Indirekte kausale Effekte

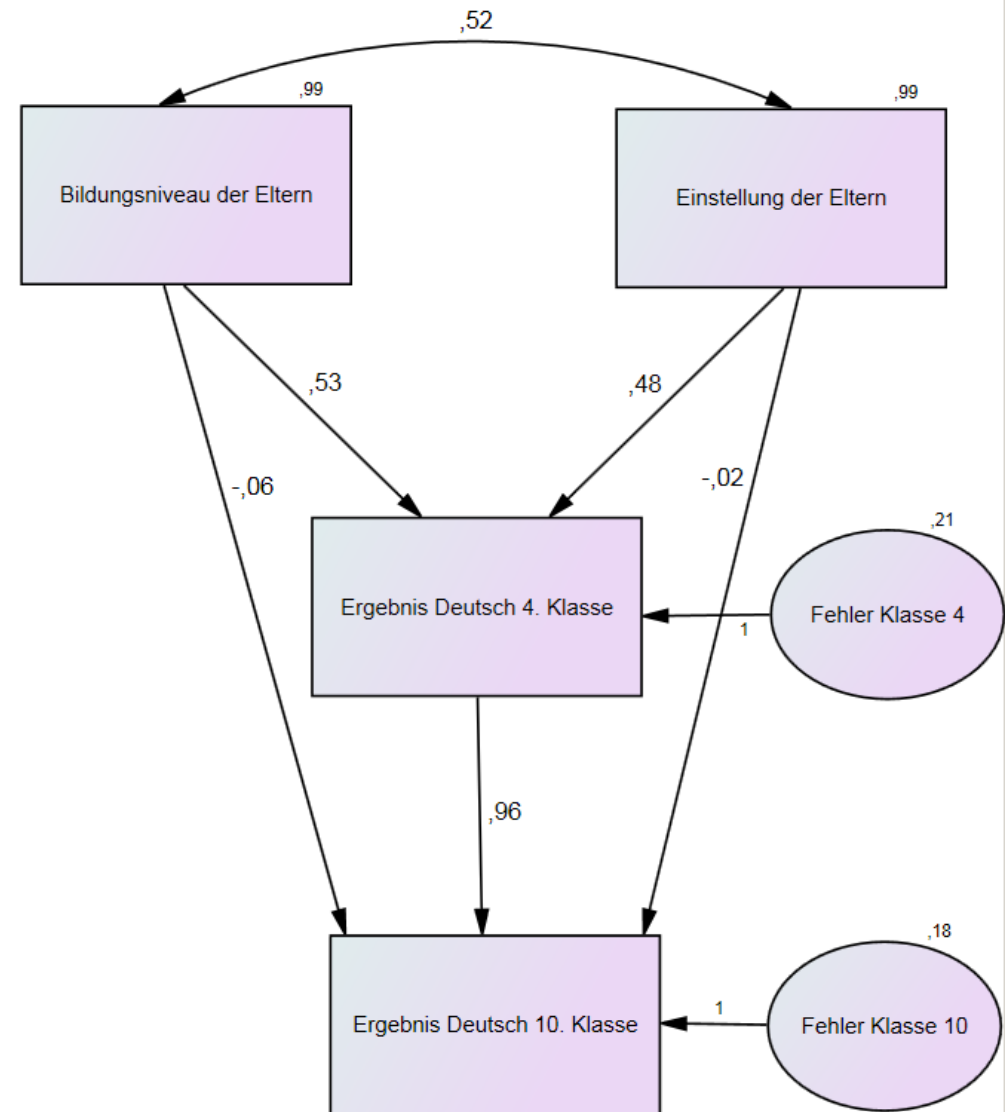
Zum Beispiel indirekter kausaler Effekt von **Bildung** auf **Deutsch 10** (Mediator: **Deutsch 4**): $0.533 \cdot 0.965 = 0.514$

Totale kausale Effekte

Zum Beispiel totaler Effekt von **Bildung** auf **Deutsch 10**:
 $-0.06 + 0.533 \cdot 0.965 = 0.454$

Indirekte korrelative Effekte

Zum Beispiel indirekter korrelativer Effekt von **Bildung** auf **Deutsch 10**:
 $0.529 \cdot 0.482 \cdot 0.965 + 0.529 \cdot -0.017 = 0.237$



➤ Pfadmodelle: Korrelationszerlegung

Fundamentaltheorem der Pfadanalyse:

Korrelationen zwischen unabhängigen (exogenen) und abhängigen (endogenen) Variablen ergeben sich als Summe der direkten und indirekten kausalen und der indirekten korrelativen Effekte.

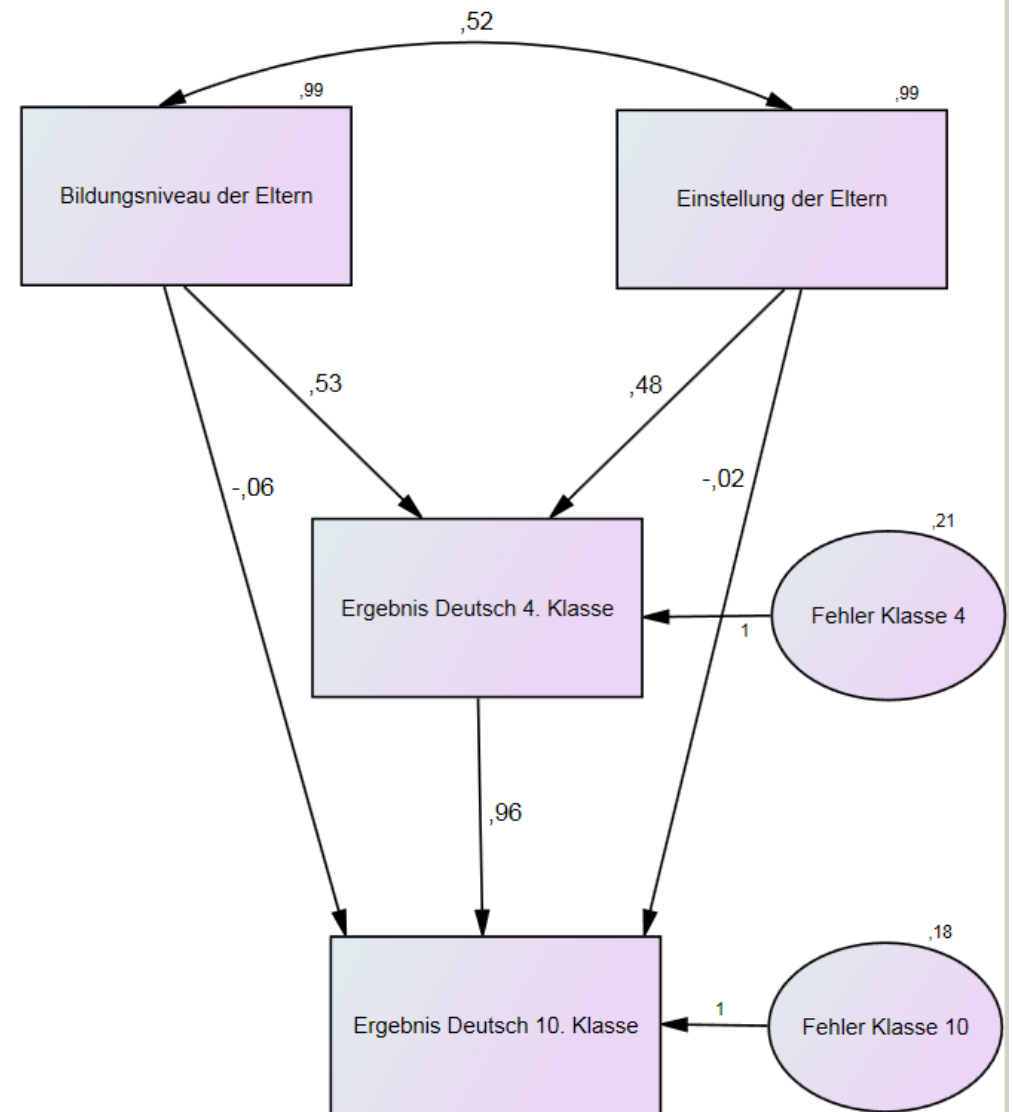
Beispiel: $r_{\text{Bildung, Deutsch 10}} = 0.691$

Direkter kausaler Effekt: -0.06

Indirekter kausaler Effekt: 0.514

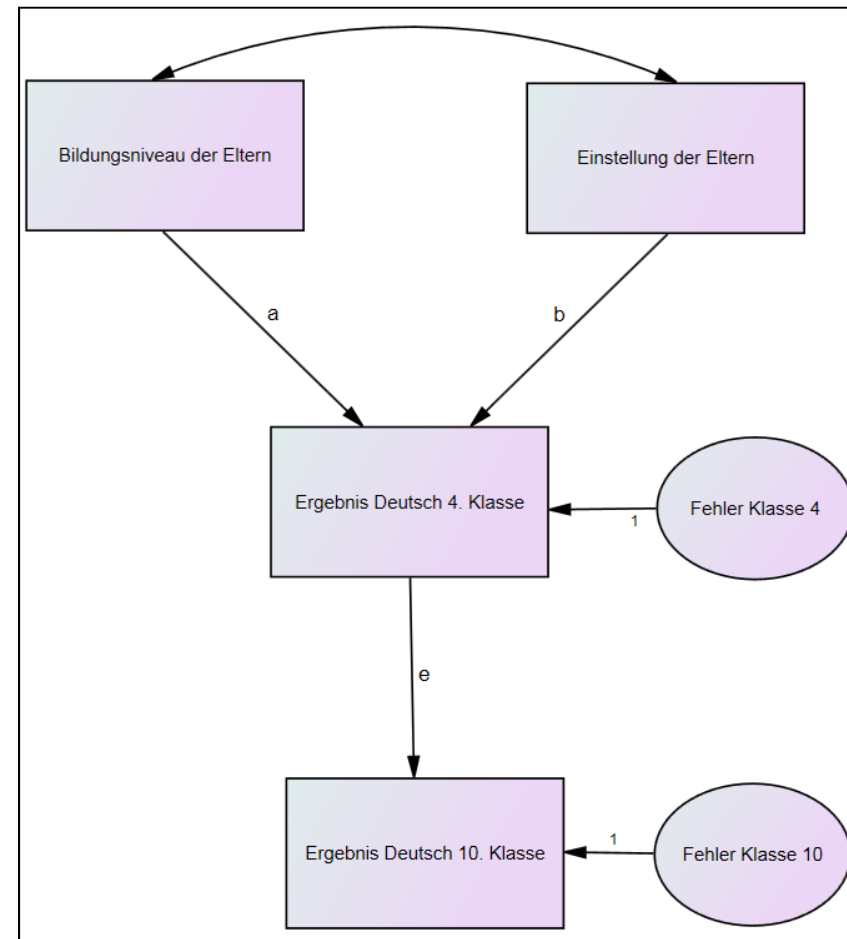
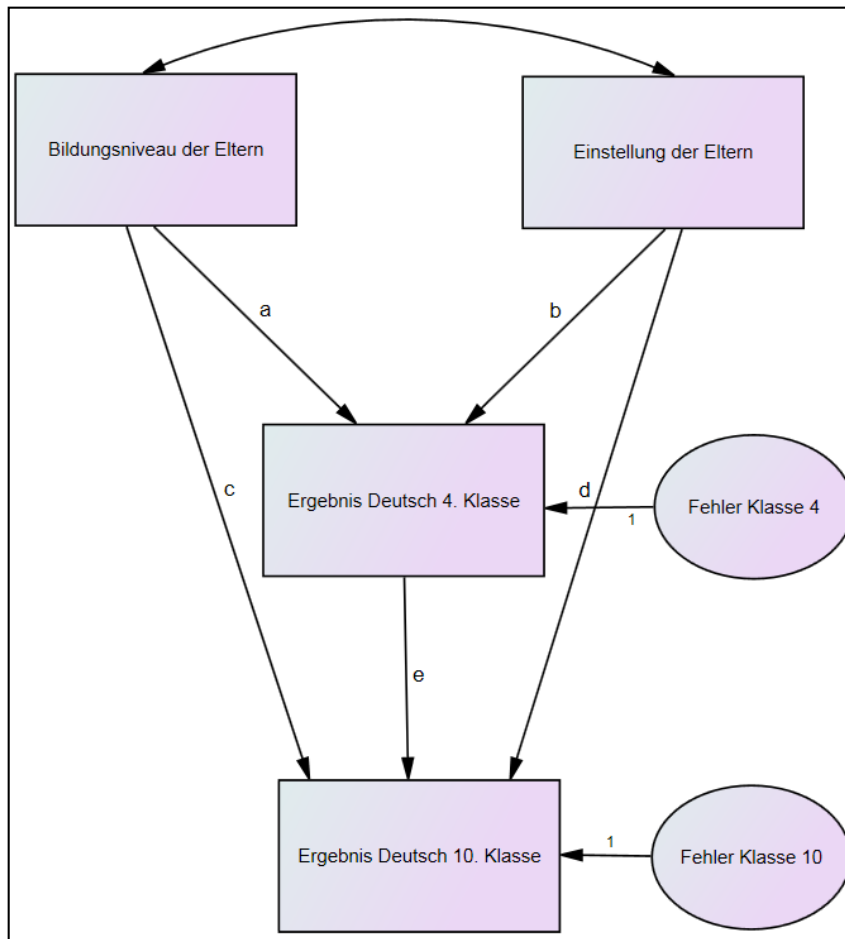
Indirekter korrelativer Effekt: 0.237

Summe der Effekte: 0.691



➤ Prinzip der Parameterschätzung

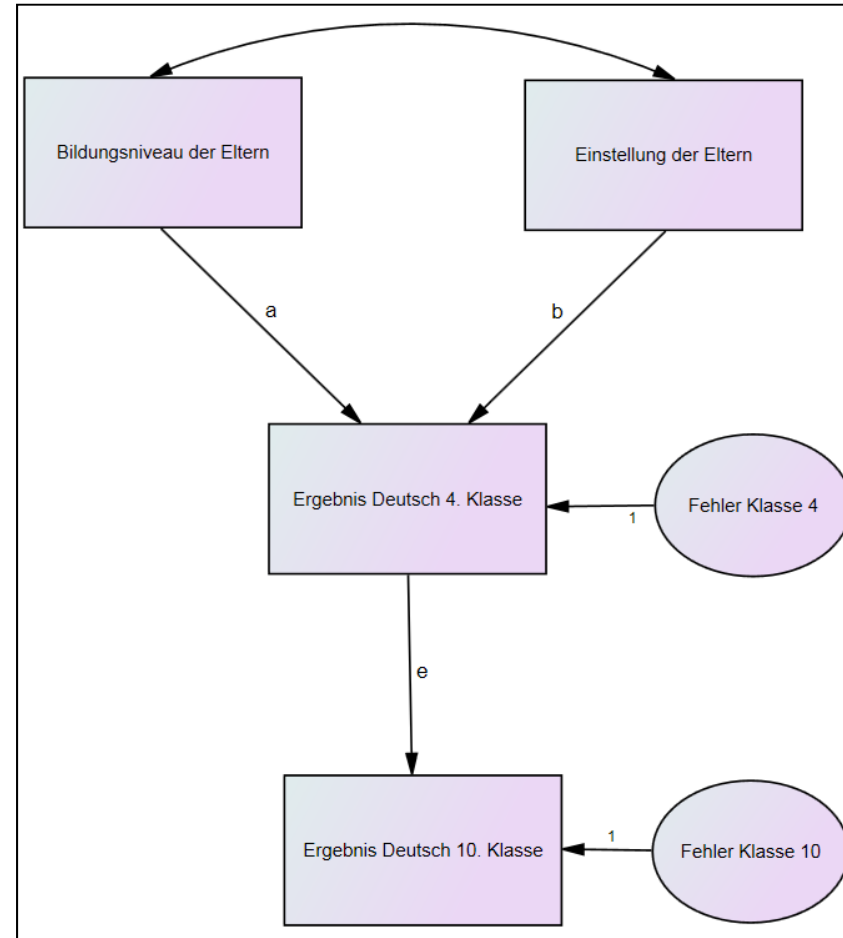
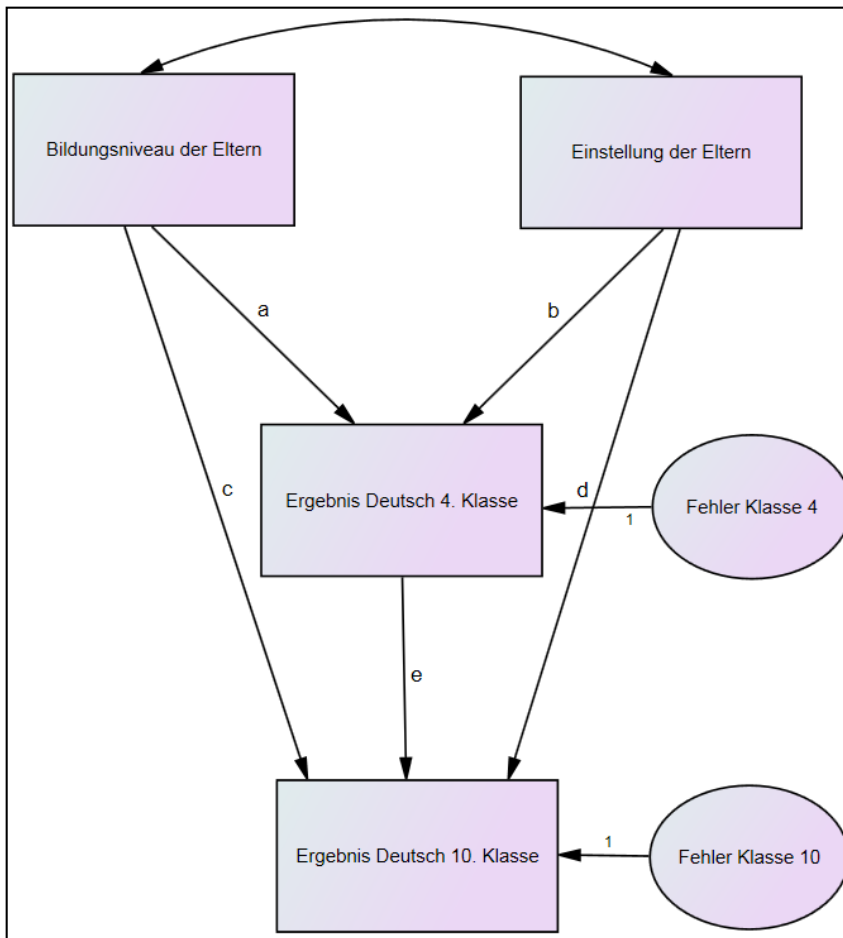
Ziel pfadanalytischer Untersuchungen: Inhaltlich aussagekräftige, „sparsame“ Modelle
Modell A: Pfade zwischen allen Variablen Modell B: Pfade entsprechend Hypothese



➤ Prinzip der Parameterschätzung

Modell A: Korrelationen zwischen Variablen können vollständig repliziert werden!

Modell B: Wie gut können Korrelationen zwischen Variablen repliziert werden?



Prinzip der Parameterschätzung

Beispiel: Wie gut könnte in Modell B die Korrelation zwischen Bildung und Deutsch 10 repliziert werden, wenn die in Modell A berechneten Pfadkoeffizienten benutzt werden würden (heuristische Betrachtung; tatsächlich werden mit der ML-Methode in Modell B abweichende Parameter geschätzt!)?

Direkter kausaler Effekt von Bildung auf Deutsch 10: 0

Indirekter kausaler Effekt von Bildung auf Deutsch 10 (Mediator: Deutsch 4):
 $0.533 \cdot 0.965 = 0.514$

Indirekter korrelativer Effekt von Bildung auf Deutsch 10:
 $0.529 \cdot 0.482 \cdot 0.965 = 0.246$

$$r_{\text{Bildung, Deutsch 10}} = 0.691$$

Direkter kausaler Effekt: 0

Indirekter kausaler Effekt: 0.514

Indirekter korrelativer Effekt: 0.246

Summe der Effekte: **0.76** → Korrelation würde bei überschätzt

Prinzip der Parameterschätzung

Allgemeines Ziel der Parameterschätzung:

Parameter so schätzen, dass durch das Modell und die für das Modell geschätzten Parameter die Beziehungen (Korrelationen bzw. Varianzen / Kovarianzen) der Variablen möglichst gut repliziert werden können.

Wichtige Methode: Maximum-Likelihood-Methode ML:

Ziel in einem iterativen Verfahren: Bestmögliche Anpassung der aus den geschätzten Parametern berechnete Varianz-/Kovarianzmatrix an die empirische Varianz-/Kovarianzmatrix; damit Maximierung der Wahrscheinlichkeit, dass die empirische Matrix auf der Grundlage der Modellparameter zustande gekommen ist.

Voraussetzung: mehrdimensionale Normalverteilung ($n > 100$).

(andere Methoden für unterschiedliche Voraussetzungen verfügbar)

➤ Prinzip der Parameterschätzung

Beispiel: Daten

Deskriptive Statistik

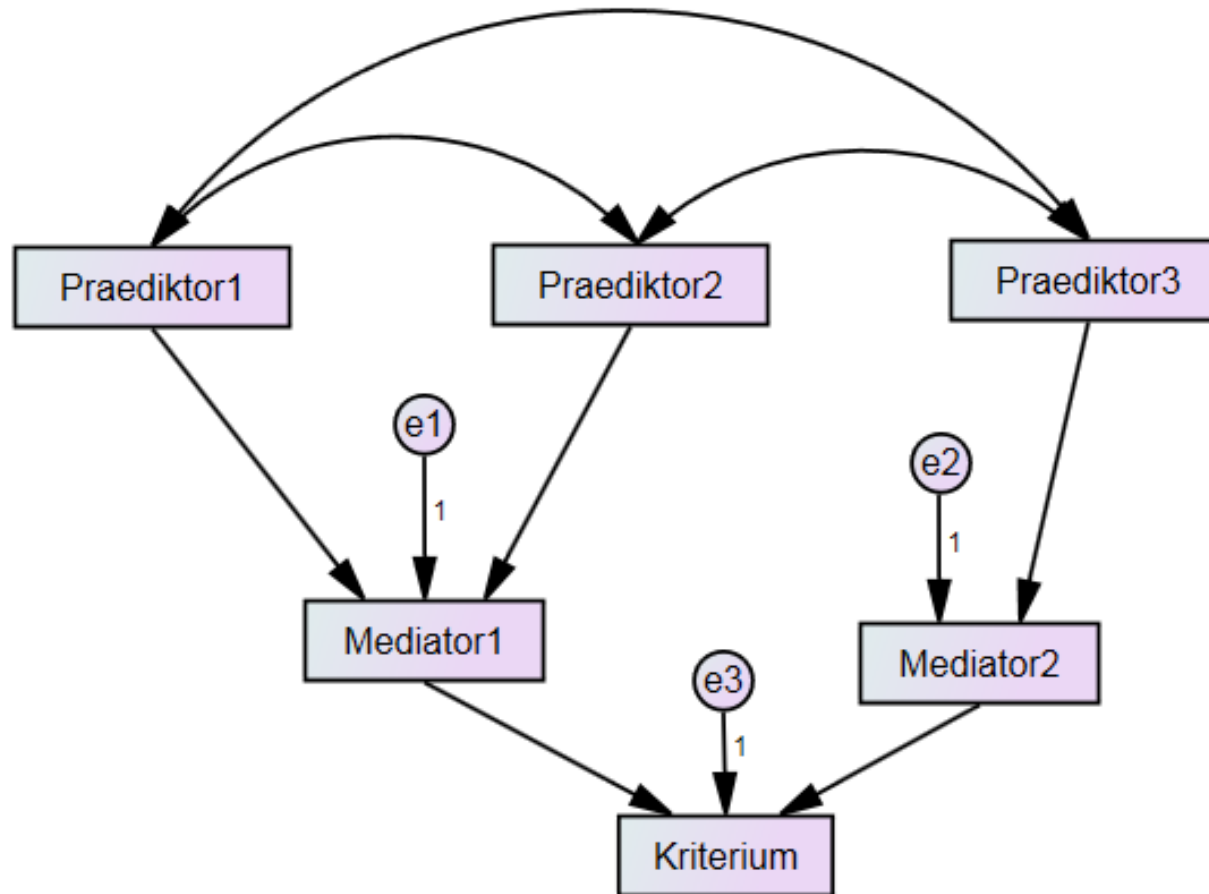
| | N | Minimum | Maximum | Mittelwert | Standardabweichung |
|--------------------------------|-----|----------|---------|------------|--------------------|
| Praediktor1 | 100 | -2,19656 | 2,45963 | ,0000000 | 1,00000000 |
| Praediktor2 | 100 | -2,91964 | 2,31116 | ,0000000 | 1,00000000 |
| Praediktor3 | 100 | -2,03713 | 2,78876 | ,0000000 | 1,00000000 |
| Mediator1 | 100 | -2,34986 | 2,37541 | ,0000000 | 1,00000000 |
| Mediator2 | 100 | -3,84415 | 2,20773 | ,0000000 | 1,00000000 |
| Kriterium | 100 | -3,23150 | 2,73790 | ,0000000 | 1,00000000 |
| Gültige Werte (Listenweise) | 100 | | | | |

Korrelationen

| | | Praediktor1 | Praediktor2 | Praediktor3 | Mediator1 | Mediator2 | Kriterium |
|-------------|--------------------------|-------------|-------------|-------------|-----------|-----------|-----------|
| Praediktor1 | Korrelation nach Pearson | 1 | -,035 | -,128 | ,568** | -,028 | ,098 |
| | Signifikanz (2-seitig) | | ,727 | ,205 | ,000 | ,780 | ,332 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Praediktor2 | Korrelation nach Pearson | -,035 | 1 | -,005 | ,632** | ,631** | ,575** |
| | Signifikanz (2-seitig) | ,727 | | ,959 | ,000 | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Praediktor3 | Korrelation nach Pearson | -,128 | -,005 | 1 | -,091 | ,449** | ,390** |
| | Signifikanz (2-seitig) | ,205 | ,959 | | ,368 | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Mediator1 | Korrelation nach Pearson | ,568** | ,632** | -,091 | 1 | ,370** | ,491** |
| | Signifikanz (2-seitig) | ,000 | ,000 | ,368 | | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Mediator2 | Korrelation nach Pearson | -,028 | ,631** | ,449** | ,370** | 1 | ,797** |
| | Signifikanz (2-seitig) | ,780 | ,000 | ,000 | ,000 | | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Kriterium | Korrelation nach Pearson | ,098 | ,575** | ,390** | ,491** | ,797** | 1 |
| | Signifikanz (2-seitig) | ,332 | ,000 | ,000 | ,000 | ,000 | |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |

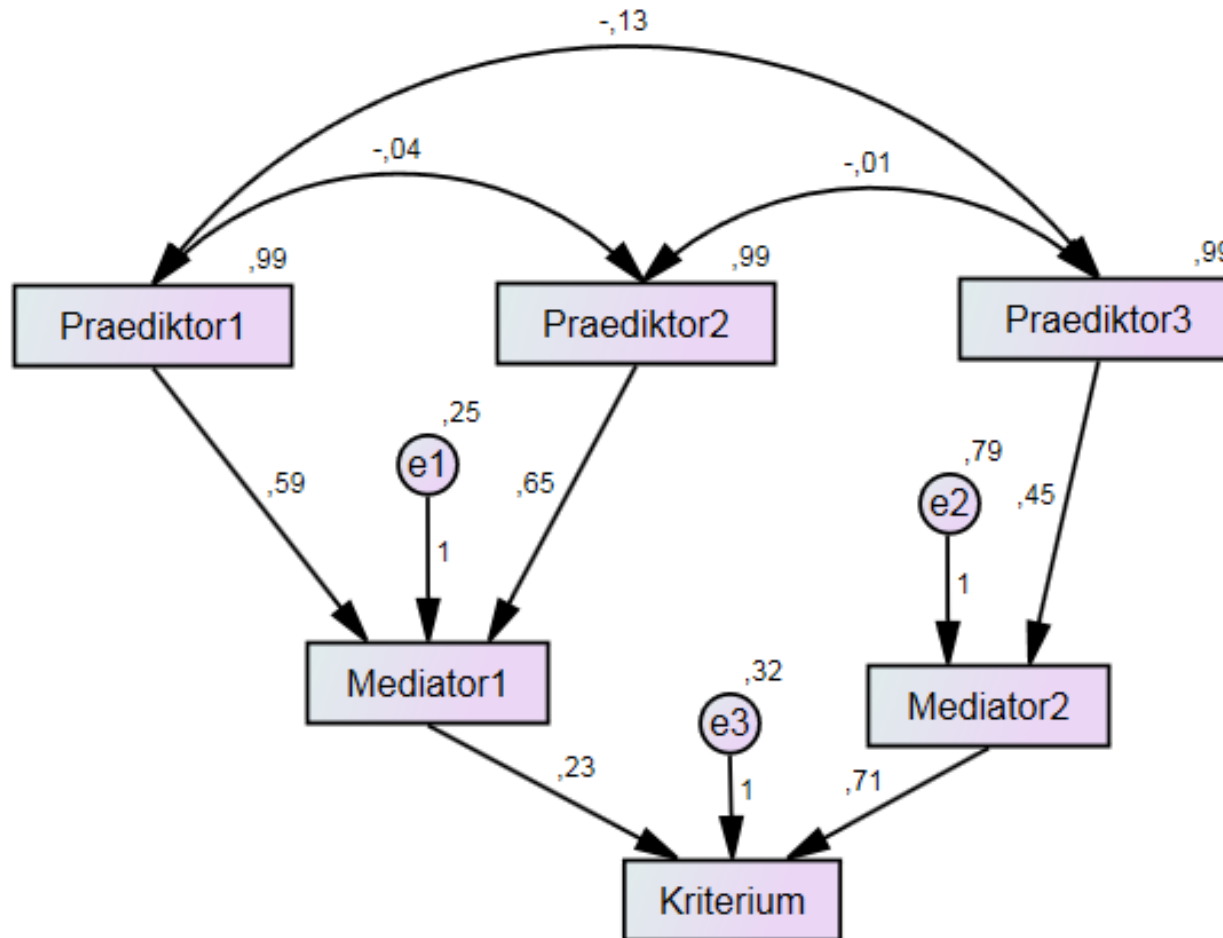
➤ Prinzip der Parameterschätzung

Beispiel: Modell



➤ Prinzip der Parameterschätzung

Beispiel: ML-Schätzungen



Prinzip der Modellgütebeurteilung

Allgemeine Frage zur Beurteilung der Modellgüte:

Wie gut ist es bei dem untersuchten Modell möglich, bei bestmöglicher Parameterschätzung die gegebene Korrelations- bzw. die Varianz-/ Kovarianzmatrix zu replizieren?

Bewertung der Gesamtstruktur: Chi-Quadrat-Test

→ Test der statistische Nullhypothese, dass die empirische Varianz-/Kovarianzmatrix der aus dem Modell ermittelten Varianz-/Kovarianzmatrix entspricht (nur bedingt anwendbar)

Bewertung der Gesamtstruktur: Fit-Indizes (unterschiedliche Indizes verfügbar)

→ zum Beispiel RMSEA (Root Mean Square Error of Approximation)

$RMSEA \leq 0.05$ → guter Modell-Fit

$RMSEA \leq 0.08$ → akzeptabler Modell-Fit

$RMSEA \geq 0.10$ → inakzeptabler Modell-Fit

➤ Prinzip der Modellgütebeurteilung

Beispiel: Modellgüte

| | |
|--|----|
| Notes for Model (Default model) | |
| Computation of degrees of freedom (Default model) | |
| Number of distinct sample moments: | 21 |
| Number of distinct parameters to be estimated: | 14 |
| Degrees of freedom (21 - 14): | 7 |
| Result (Default model) | |
| Minimum was achieved | |
| Chi-square = 73,890 | |
| Degrees of freedom = 7 | |
| Probability level = ,000 | |

RMSEA = 0.31

Hypothesenüberprüfung

Bewertung von Teilstrukturen

→ Test einzelner Modellparameter: Tests auf der Grundlage der Schätzwerte und der Standardfehler

→ Tests indirekter Effekte: Bootstrap-Tests

→ Residuenanalyse

→ ggf. Vergleich alternativer Schätzverfahren

Modifikationen des Modells

Bei unbefriedigender Modellgüte

- Modifikation der Modellstruktur auf der Grundlage der Daten und inhaltlicher Überlegungen prüfen
- Exploratorischer (!!) Analyseschritt
- Prüfung neuer Modelle mit neuen Daten erforderlich!

Statistische Entscheidungshilfen zur Modellmodifikation:

- Entscheidung über Hinzunahme von Pfaden, Kovarianzen u.a. auf der Basis von Modifikations-Indizes: Angabe, um welchen Betrag sich der Chi-Quadrat-Wert verändert (d.h. wie stark die Modellgüte verbessert wird), wenn ein zusätzlicher zu schätzender Parameter im Modell frei gegeben wird.
- Entscheidung über Entfernen von Pfaden, Kovarianzen u.a. auf der Basis der Ergebnisse der statistischen Tests der jeweiligen Koeffizienten (ggf. Entfernen „nicht signifikanter“ Pfade).
- Entscheidungen immer unter vorrangiger Einbeziehung inhaltlicher Gesichtspunkte treffen

➤ Modifikationen des Modells

Beispiel: Modification Indices

The screenshot shows the Amos Output window for a model named 'Pfadmodell_Beispiel_1.amw'. The 'Modification Indices' section is highlighted in the left-hand navigation pane. The main content area displays the following sections:

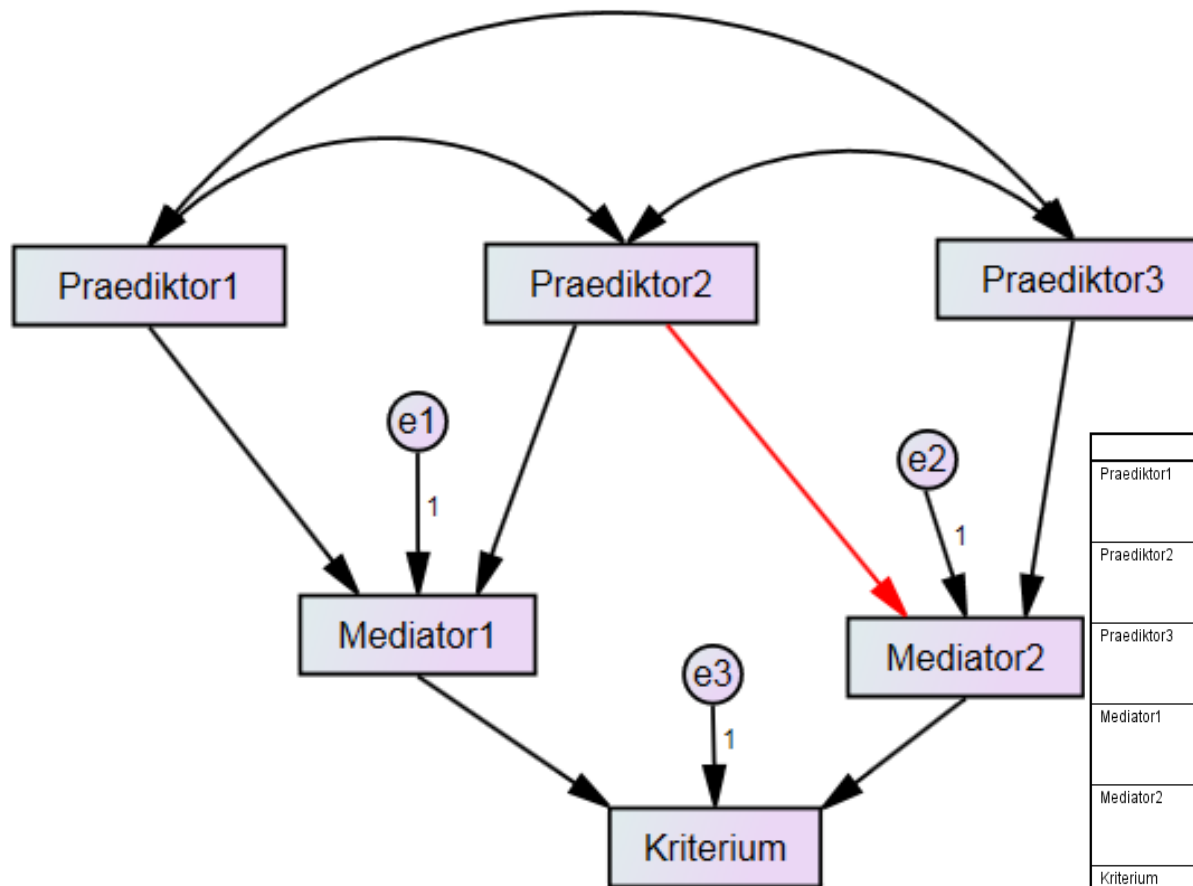
- Modification Indices (Group number 1 - Default model)**
- Covariances: (Group number 1 - Default model)**
- Variances: (Group number 1 - Default model)**
- Regression Weights: (Group number 1 - Default model)**

| | M.I. | Par Change |
|----------------------|--------|------------|
| e2 <--> Praedikator2 | 49,923 | ,628 |

| | M.I. | Par Change |
|-----------------------------|--------|------------|
| Mediator2 <--- Praedikator2 | 49,688 | ,633 |
| Mediator2 <--- Mediator1 | 20,972 | ,411 |

➤ Modifikationen des Modells

Beispiel: Modifiziertes Modell

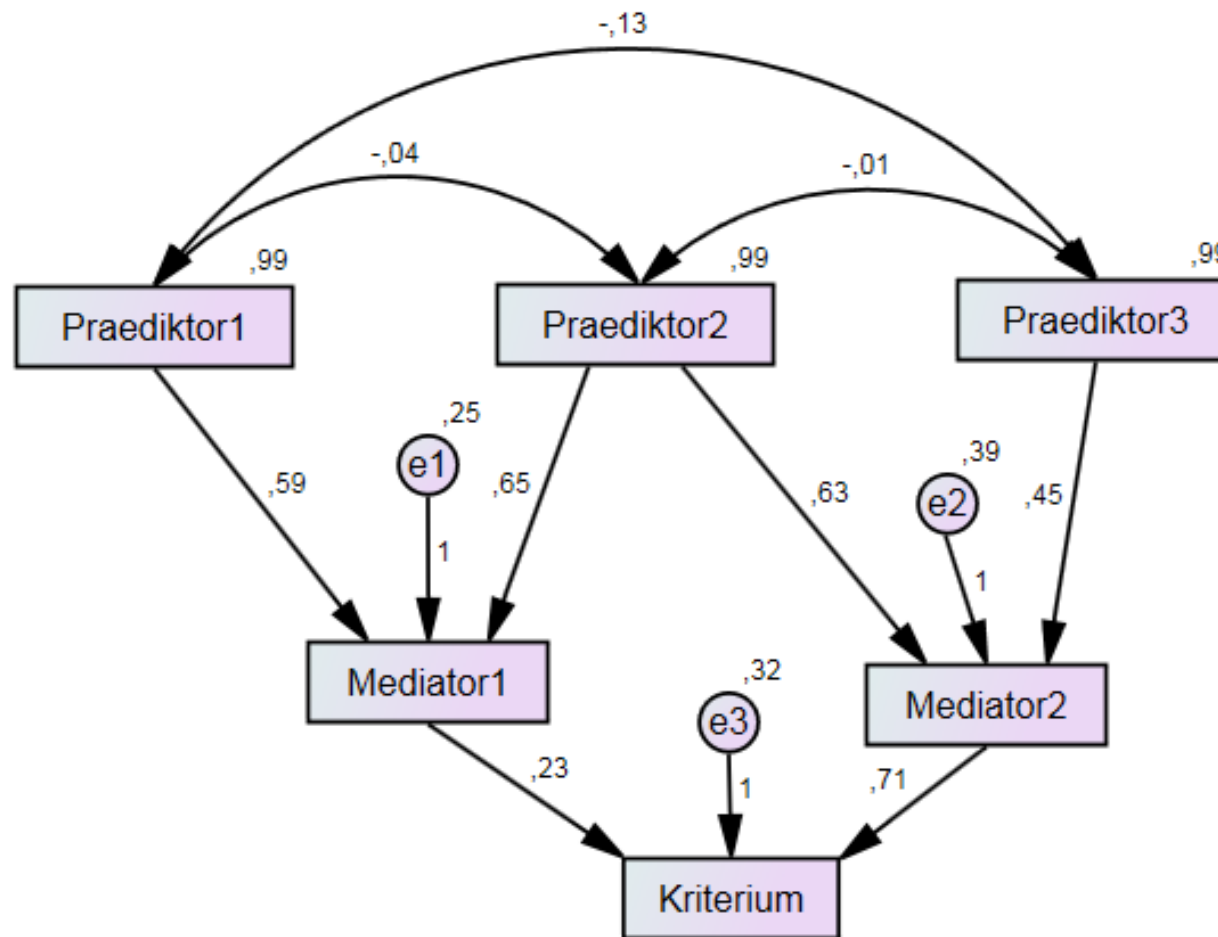


Korrelationen

| | | Praediktor1 | Praediktor2 | Praediktor3 | Mediator1 | Mediator2 | Kriterium |
|-------------|--------------------------|-------------|-------------|-------------|-----------|-----------|-----------|
| Praediktor1 | Korrelation nach Pearson | 1 | -,035 | -,128 | ,568** | -,028 | ,098 |
| | Signifikanz (2-seitig) | | ,727 | ,205 | ,000 | ,780 | ,332 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Praediktor2 | Korrelation nach Pearson | -,035 | 1 | -,005 | ,632** | ,631** | ,575** |
| | Signifikanz (2-seitig) | ,727 | | ,959 | ,000 | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Praediktor3 | Korrelation nach Pearson | -,128 | -,005 | 1 | -,091 | ,449** | ,390** |
| | Signifikanz (2-seitig) | ,205 | ,959 | | ,368 | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Mediator1 | Korrelation nach Pearson | ,568** | ,632** | -,091 | 1 | ,370** | ,491** |
| | Signifikanz (2-seitig) | ,000 | ,000 | ,368 | | ,000 | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Mediator2 | Korrelation nach Pearson | -,028 | ,631** | ,449** | ,370** | 1 | ,797** |
| | Signifikanz (2-seitig) | ,780 | ,000 | ,000 | ,000 | | ,000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |
| Kriterium | Korrelation nach Pearson | ,098 | ,575** | ,390** | ,491** | ,797** | 1 |
| | Signifikanz (2-seitig) | ,332 | ,000 | ,000 | ,000 | ,000 | |
| | N | 100 | 100 | 100 | 100 | 100 | 100 |

➤ Modifikationen des Modells

Beispiel: ML-Schätzungen im modifizierten Modell



➤ Modifikationen des Modells

Beispiel: Modellgüte im modifizierten Modell

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

| | |
|--|----|
| Number of distinct sample moments: | 21 |
| Number of distinct parameters to be estimated: | 15 |
| Degrees of freedom (21 - 15): | 6 |

Result (Default model)

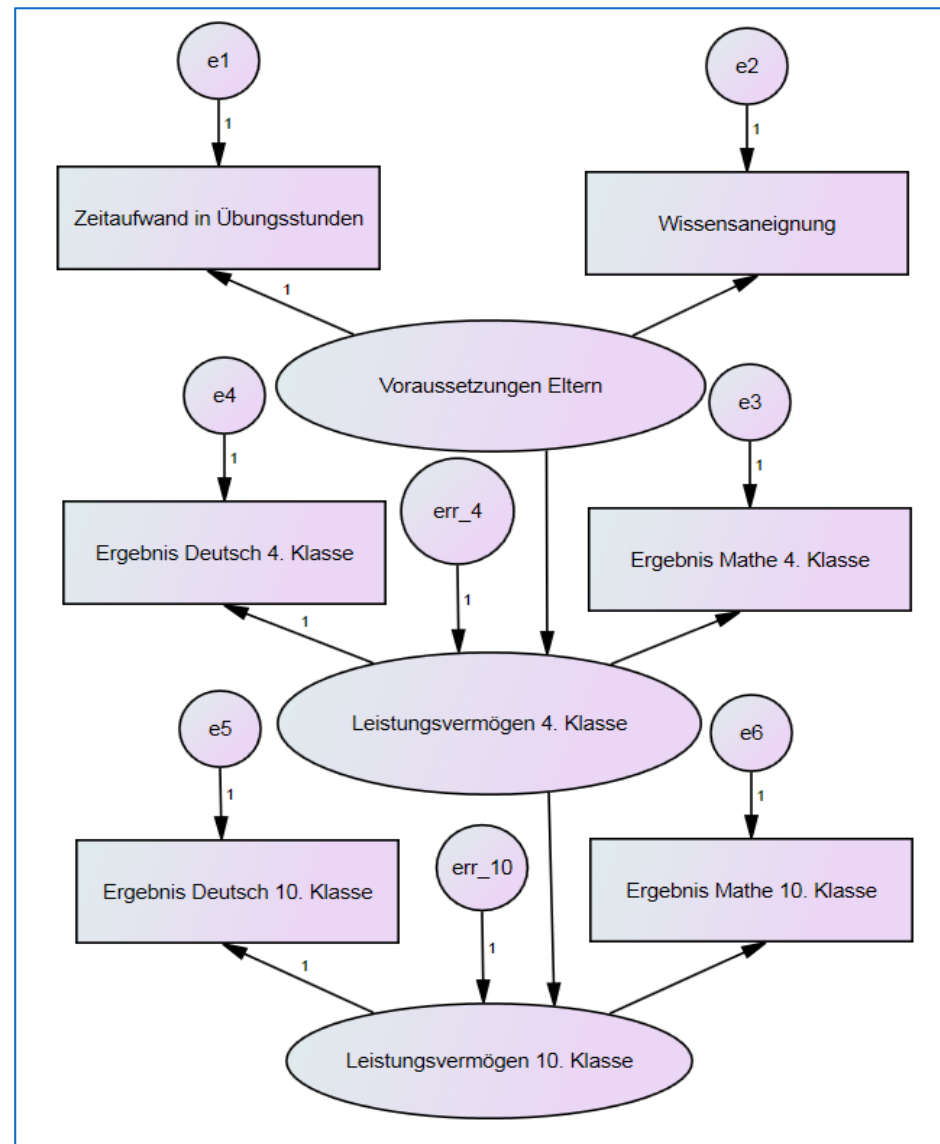
Minimum was achieved
Chi-square = 4,889
Degrees of freedom = 6
Probability level = ,558

RMSEA = 0.00

➤ LSM mit latenten Variablen

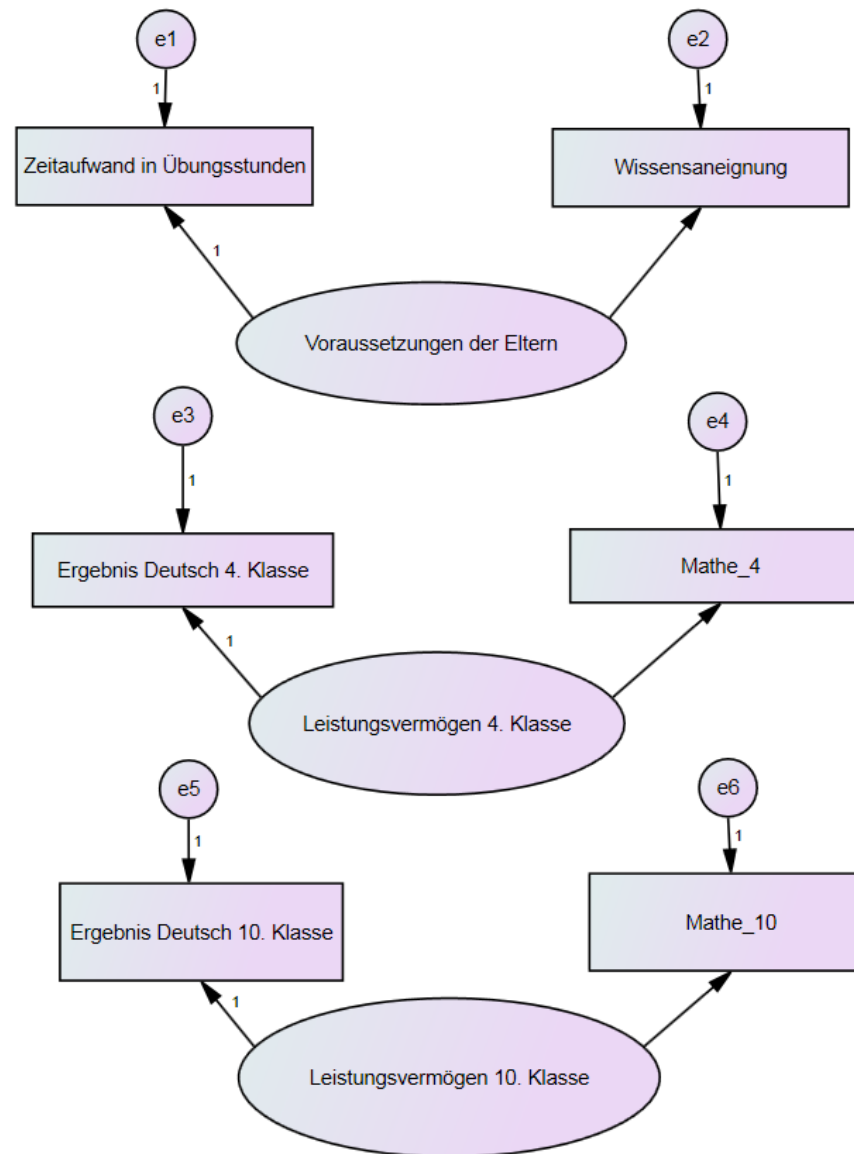
Verallgemeinerung der Pfadanalyse, Verbindung mit der (konfirmatorischen) Faktorenanalyse

→ Pfadmodelle mit latenten Variablen



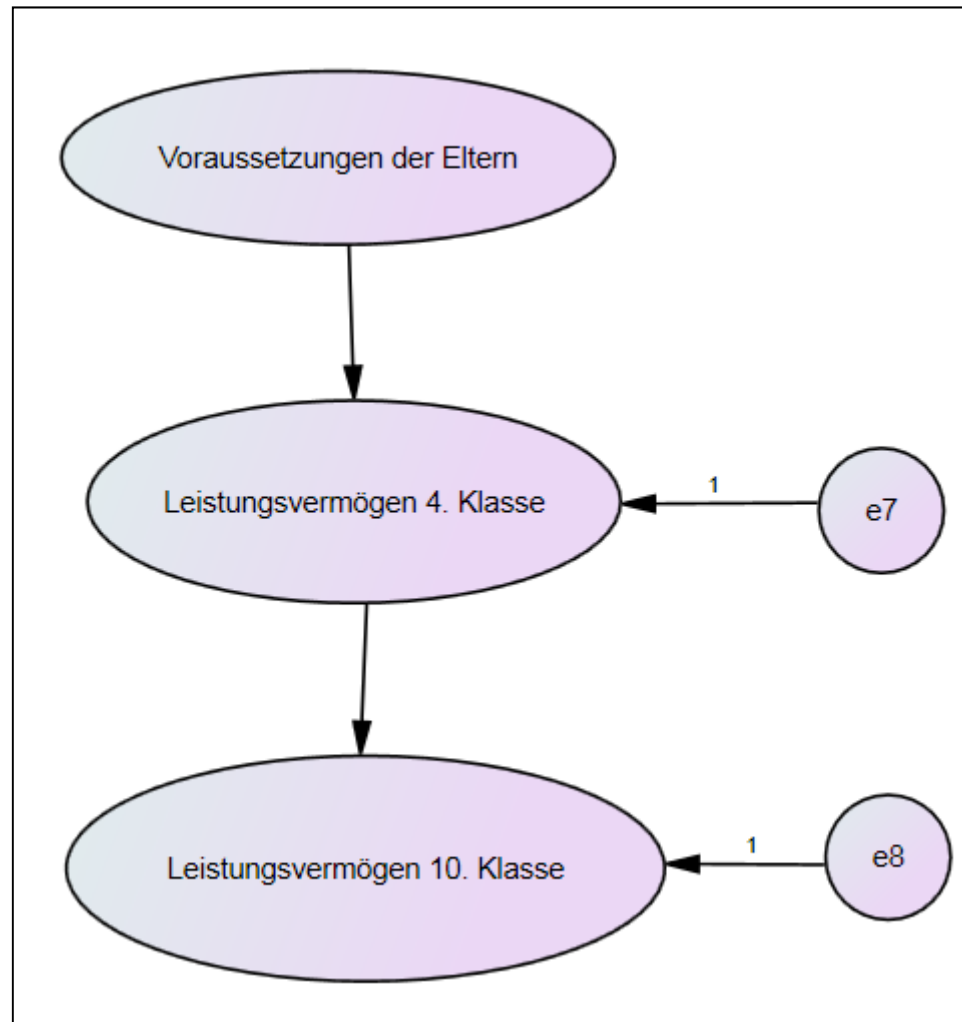
Messmodell

Messmodelle beschreiben, welche manifesten Variablen Indikatoren von latenten Variablen sind.



➤ Strukturmodell

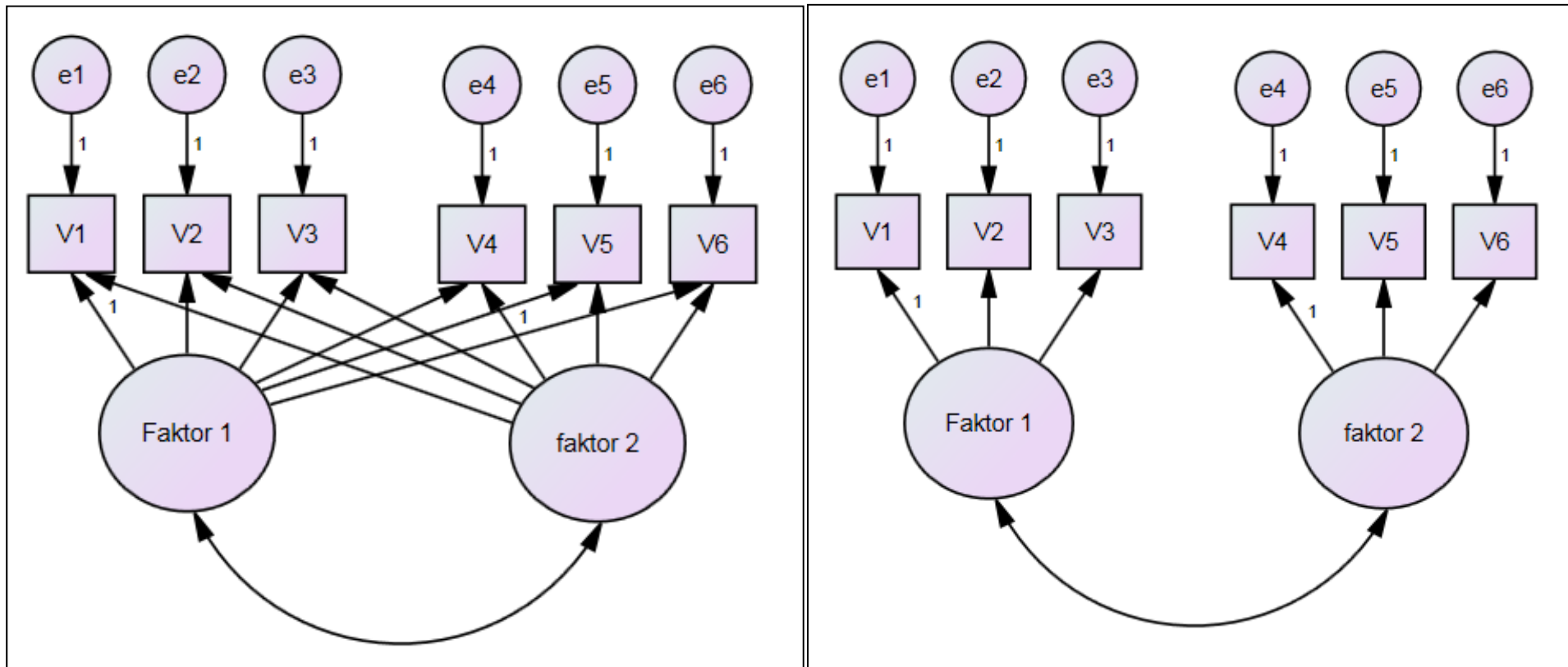
Strukturmodelle beschreiben die Beziehungen der latenten Variablen im untersuchten Modell.



➤ Konfirmatorische Faktorenanalyse

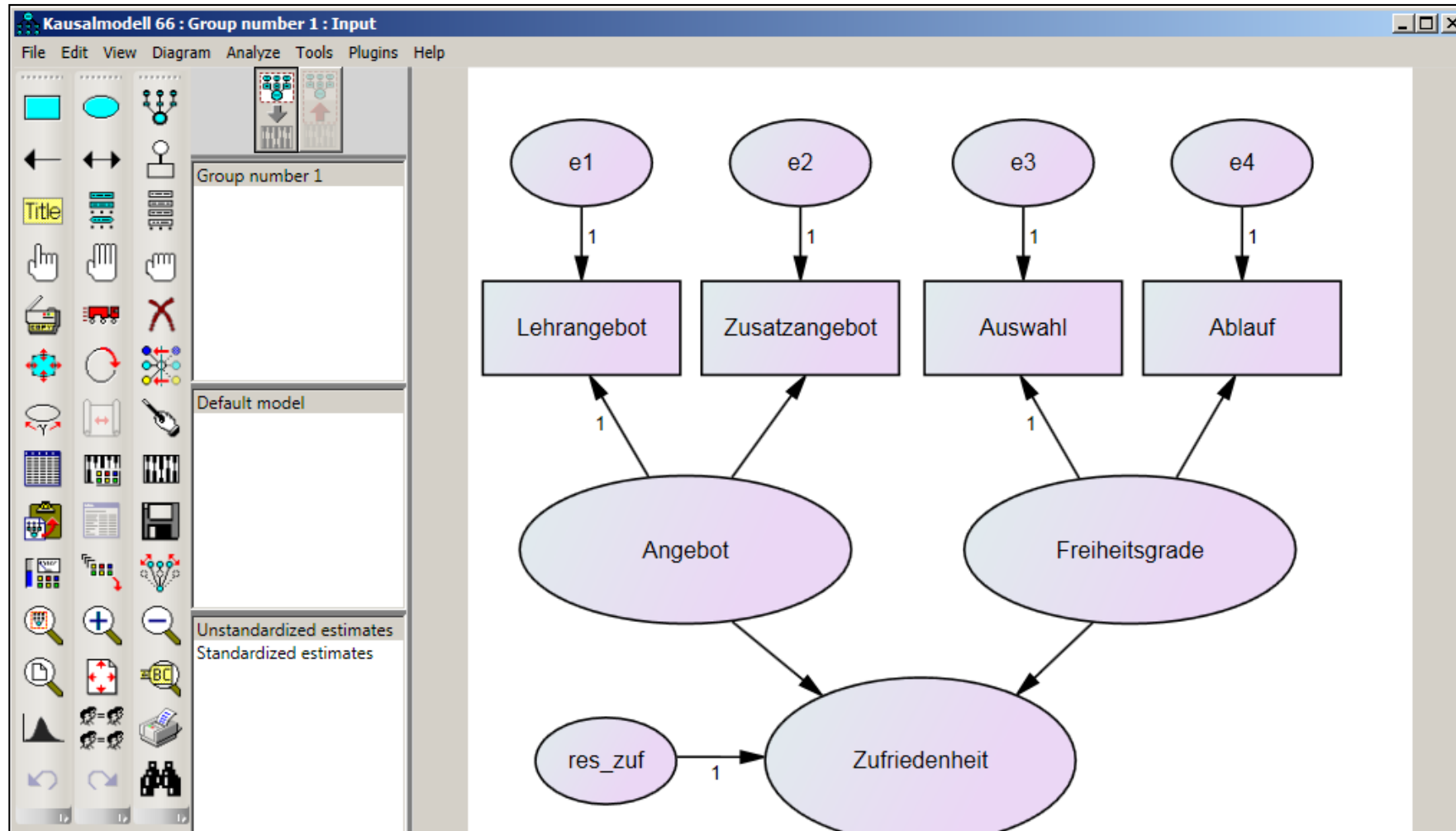
Exploratorische Faktorenanalyse

Konfirmatorische Faktorenanalyse



(Abbildung nach Bühner, 2006, S. 260)

Softwarebeispiel: IBM SPSS AMOS



Einige ausgewählte Literaturhinweise

- Arbuckle, J. L. & Wothke, W. (1999). *Amos 4.0 User's Guide*. Chicago, IL.: Small Waters Corporation.
- Blunch, N.J. (2013). *Introduction to Structural Equation Modeling Using IBM SPSS and AMOS*(2. ed.). London: Sage.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Aufl.). München: Pearson-Studium.
- Byrne, B. M. (2009). *Structural Equation Modelling With Amos: Basic Concepts, Applications, and Programming* (2. ed.). London: Chapman & Hall/CRC.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden*. Weinheim: Beltz.
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (Methodology in the Social Sciences) (3rd ed.). New York: Guilford Press.
- Rudolf, M. & Müller, J. (2012). *Multivariate Verfahren. Eine praxisorientierte Einführung mit Anwendungsbeispielen in SPSS* (2. Aufl.). Göttingen: Hogrefe.
- Schumacker, R. E. & Lomax, R. G. (2009). *A Beginner's Guide to Structural Equation Modelling*. London: Taylor & Francis.
- Weiber, R. & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* (2. Aufl.). Heidelberg: Springer.