Das allgemeine Strukturgleichungsmodell mit latenten Variablen

Jost Reinecke

Universität Bielefeld

4. Januar 2006

Die Verbindung von Pfad- und Faktorenanalyse

- Die Modellspezifikation
- Die Schätzung der Modellparameter
- Die Identifikation der Modellparameter
- Standardisierte und unstandardisierte Koeffizienten
- Die Effektzerlegung
- Multiple Gruppenvergleiche
- Empirische Beispiele

Kategoriales Meßniveau in Strukturgleichungsmodellen

Strukturgleichungsmodell als Markov-Modell

Strukturgleichungsmodell als Wachstumsmodell

Die Verbindung von Pfad- und Faktorenanalyse

Die Variablenbezeichnungen

Zeichen	Aussprache	Bedeutung
ξ	Ksi	exogene latente Variable
η	Eta	endogene latente Variable
x		manifeste Variable der latenten Variablen ξ
y		manifeste Variable der latenten Variablen η
ζ	Zeta	Residuen der latenten Variablen η
δ	Delta	Meßfehler der manifesten Variablen x
ϵ	Epsilon	Meßfehler der manifesten Variablen y

Die Matrizen des Meßmodells

Zeichen	Aussprache	Bedeutung
Λ_x	Lambda$_x$	$(p \times m)$ Faktorenladungen der x-Variablen
Λ_y	Lambda$_y$	$(q \times n)$ Faktorenladungen der y-Variablen
Θ_δ	Theta-delta	$(p \times p)$ Meßfehlervarianzen und -kovarianzen der x-Variablen
Θ_ϵ	Theta-epsilon	$(q \times q)$ Meßfehlervarianzen und -kovarianzen der y-Variablen
Φ	Phi	$(n \times n)$ Varianzen und Kovarianzen zwischen den ξ-Variablen

Die Parameter des Strukturmodells

Zeichen	Aussprache	Bedeutung
γ	gamma	Element der Matrix Γ (Strukturkoeffizient)
β	beta	Element der Matrix B (Strukturkoeffizient)
ϕ	phi	Element der Matrix Φ (Varianz/Kovarianz)
ψ	psi	Element der Matrix Ψ (Varianz/Kovarianz)

Die Modellspezifikation

Die Modellierung von Kausalbeziehungen zwischen den abhängigen latenten Variablen (Faktoren erster Ordnung) erlaubt das allgemeine Strukturgleichungsmodell. Diese werden in der folgenden Gleichung über die Matrix B spezifiziert:

$$\eta = B\eta + \Gamma\xi + \zeta \tag{1}$$

η ist der Vektor der *abhängigen* latenten Variablen, ξ der Vektor der *unabhängigen* latenten Variablen. In der Matrix B werden die Beziehungen der *abhängigen* latenten Variablen untereinander spezifiziert, in der Matrix Γ die Beziehungen zwischen *unabhängigen* und *abhängigen* latenten Variablen. Vektor ζ bezeichnet die Residuen der *abhängigen* latenten Variablen, die annahmegemäß einen Erwartungswert von Null haben $(\mathrm{E}(\zeta)) = 0)$. Die Meßfehlervarianzen σ_ζ werden in der Matrix Ψ spezifiziert. Desweiteren wird davon ausgegangen, daß die Differenz $I - B$ nicht singulär ist. Die Varianzen und Kovarianzen der unabhängigen latenten Variablen ξ werden in der Matrix Φ angegeben.

Neben der Strukturebene ist das Meßmodell zu spezifizieren:

$$x = \Lambda_x \xi + \delta \tag{2}$$

$$y = \Lambda_y \eta + \epsilon \tag{3}$$

x ist der Vektor der Indikatoren für die latenten Variablen η, y ist der Vektor der Indikatoren für die latenten Variablen ξ. Die Beziehungen zwischen den latenten und manifesten Variablen werden durch die Koeffizientenmatrizen Λ_x und Λ_y ausgedrückt. Die Vektoren δ und ϵ beinhalten die Meßfehleranteile der manifesten Variablen. Die Meßfehler haben nach den üblichen Annahmen einen Erwartungswert von Null ($E(\delta) = E(\epsilon) = 0$) und sind mit den jeweiligen latenten Variablen unkorreliert ($E(\xi\delta) = E(\eta\epsilon) = 0$). Die Meßfehlervarianzen σ_δ und σ_ϵ werden in den Matrizen Θ_δ und Θ_ϵ spezifiziert.

Folgende Spezialfälle subsumieren sich unter das in den Gleichungen 1 bis 3 spezifizierte allgemeine Strukturgleichungsmodell:

- Wenn kein Meßmodell existiert ($\Lambda_x = I$, $\Lambda_y = I$, $\Theta_\delta = 0$, $\Theta_\epsilon = 0$), dann reduziert sich die Strukturgleichung 1 auf

$$y = By + \Gamma x + \zeta \tag{4}$$

  Diese Formalisierung entspricht einem Pfadmodell mit ausschließlich gemessenen Variablen.

- Wird für die unabhängigen latenten Variablen kein Meßmodell spezifiziert ($\Lambda_x = I$, $\Theta_\delta = 0$) und existiert nur eine abhängige latente Variable η_1, dann reduziert sich die Strukturgleichung 1 auf

$$\eta_1 = \Gamma x + \zeta_1 \tag{5}$$

  und die Meßgleichung 3 auf

$$x = \xi \tag{6}$$

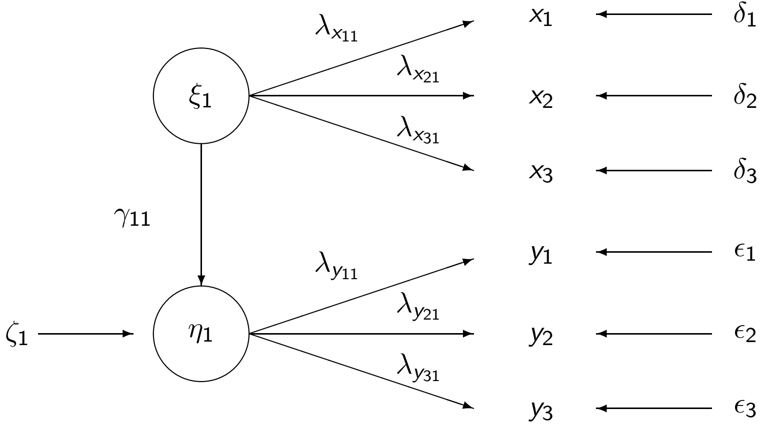
Die Meßgleichung 2 wird dann nur für eine latente Variable η_1 formuliert:

$$y = \Lambda_y \eta_1 + \epsilon \tag{7}$$

Diese Formalisierung entspricht einem sogenannten *Multiple-Indicator-Multiple-Cause*-Modell.

- Wird nur ein Meßmodell spezifiziert (Gleichung 2) sowie die Parameter der Kovarianz- bzw. Korrelationsmatrix Φ, dann reduziert sich das Strukturgleichungsmodell auf ein konfirmatorisches Faktorenmodell.
- Wird ein Faktorenmodell zweiter Ordnung spezifiziert, dann ist Matrix $B = 0$ und es existiert nur ein Meßmodell für die y-Indikatoren.

Strukturgleichungsmodell mit zwei latenten Variablen



Den allgemeinen Struktur- und Meßgleichungen folgend, läßt sich das Modell aus Abbildung 1 formal spezifizieren:

$$\eta_1 = \gamma_{11}\xi_1 + \zeta_1 \tag{8}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda_{x_{11}} \\ \lambda_{x_{21}} \\ \lambda_{x_{31}} \end{pmatrix} * \begin{pmatrix} \xi_1 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} \tag{9}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \lambda_{y_{11}} \\ \lambda_{y_{21}} \\ \lambda_{y_{31}} \end{pmatrix} * \begin{pmatrix} \eta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \tag{10}$$

mit den Annahmen

$$\begin{aligned} COV(\xi_i, \delta_j) &= 0 \\ COV(\eta_i, \epsilon_j) &= 0 \end{aligned}$$

für alle i und j und

$$\begin{aligned} E(\delta_j) &= 0 \\ E(\epsilon_j) &= 0 \end{aligned}$$

für alle j.

Die Spalten der Matrizen Λ_x und Λ_y korrespondieren zu den latenten Variablen ξ_1 bzw. η_1. Die Parameter in den Matrizen zeigen an, welche manifeste Variable zur Messung der latenten Variable herangezogen wird. Der Vektor δ enthält die Meßfehler für die manifesten Variablen x, der Vektor ϵ die Meßfehler für die manifesten Variablen y. Während bei den dargestellten Spezialfällen nur einige der Parametermatrizen verwendet werden, sind bei einem allgemeinen Strukturgleichungsmodell alle acht Parametermatrizen zu spezifizieren. Für das Modell in Abbildung 1 werden diese im folgenden aufgelistet:

$$\Lambda_x = \begin{pmatrix} \lambda_{x_{11}} \\ \lambda_{x_{21}} \\ \lambda_{x_{31}} \end{pmatrix} \qquad \Lambda_y = \begin{pmatrix} \lambda_{y_{11}} \\ \lambda_{y_{21}} \\ \lambda_{y_{31}} \end{pmatrix} \tag{11}$$

$$\Theta_\delta = \begin{pmatrix} \sigma_{\delta_1} & & \\ 0 & \sigma_{\delta_2} & \\ 0 & 0 & \sigma_{\delta_3} \end{pmatrix} \qquad \Theta_\epsilon = \begin{pmatrix} \sigma_{\epsilon_1} & & \\ 0 & \sigma_{\epsilon_2} & \\ 0 & 0 & \sigma_{\epsilon_3} \end{pmatrix}$$

$$B = \begin{pmatrix} 0 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} \gamma_{11} \end{pmatrix} \qquad \Phi = \begin{pmatrix} \phi_{11} \end{pmatrix} \qquad \Psi = \begin{pmatrix} \psi_{11} \end{pmatrix}$$

Die Schätzung der Modellparameter

Alle zu schätzenden Parameter werden im Parametervektor Θ zusammengefaßt. Die geschätzte Kovarianzmatrix Σ ist eine Funktion des Parametervektors Θ:

$$\Sigma(\Theta) = \left[\begin{array}{cc} \Sigma_{yy}(\Theta) & \Sigma_{yx}(\Theta) \\ \Sigma_{xy}(\Theta) & \Sigma_{xx}(\Theta) \end{array} \right] \tag{12}$$

Im folgenden werden die vier Elemente der Matrix $\Sigma(\Theta)$ getrennt hergeleitet. $\Sigma_{yy}(\Theta)$ enthält die Kovarianzen der y-Variablen als Funktion des Parametervektors Θ und ist gleich dem Erwartungswert des Produktes yy':

$$\begin{aligned} \Sigma_{yy}(\Theta) &= E(yy') \\ &= E[(\Lambda_y \eta + \epsilon)(\eta' \Lambda_y' + \epsilon')] \\ &= \Lambda_y E(\eta\eta')\Lambda_y' + \Theta_\epsilon \end{aligned} \tag{13}$$

Der Term $E(\eta\eta')$ kann durch Substitution mit Gleichung 1 weiter zerlegt werden. So wird erkennbar, daß die Kovarianz der y-Indikatoren in eine komplexe Funktion aus sechs der acht oben angegebenen Parametermatrizen (vgl. die Spezifikation in 11) zerlegt werden kann:

$$\Sigma_{yy}(\Theta) = \Lambda_y(I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)[(I-B)^{-1})'\Lambda_y' + \Theta_\epsilon \tag{14}$$

Der rechte obere Quadrant der Matrix $\Sigma(\Theta)$ enthält die Kovarianzen der y- und x-Variablen:

$$\begin{aligned} \Sigma_{yx}(\Theta) &= E(yx') \\ &= E[(\Lambda_y\eta + \epsilon)(\xi'\Lambda_x' + \delta')] \\ &= \Lambda_y E(\eta\xi')\Lambda_x' \end{aligned} \tag{15}$$

Wird wiederum durch $\eta = (I-B)^{-1}(\Gamma\xi + \zeta)$ substituiert, dann erhält man:

$$\Sigma_{yx}(\Theta) = \Lambda_y(I-B)^{-1}\Gamma\Phi\Lambda_x' \tag{16}$$

Der linke untere Quadrant der Matrix $\Sigma(\Theta)$ ist die Transponierte von $\Sigma_{yx}(\Theta)$:

$$\Sigma_{xy}(\Theta) = \Lambda_x\Phi\Gamma'[(I-B)^{-1}]'\Lambda_y' \tag{17}$$

Als letztes sind die Kovarianzen der x-Variablen (Σ_{xx}) abzuleiten, was im Rahmen der Meßmodelle schon behandelt wurde:

$$\begin{aligned} \Sigma(\Theta) &= E(xx') \\ &= E[(\lambda_x \xi + \delta)(\lambda_x' \xi' + \delta')] \\ &= \lambda_x E(\xi\xi')\lambda_x' + \Theta_\delta \\ &= \lambda_x \Phi \lambda_x' + \Theta_\delta \end{aligned} \tag{18}$$

Wenn Gleichungen 14, 16, 17 und 18 in Gleichung 12 eingesetzt werden, erhält man die geschätzte Varianz-/Kovarianzmatrix Σ als Funktion des Parametervektors Θ:

$$\begin{aligned} \Sigma(\Theta) &= \left[\begin{array}{cc} \Sigma_{yy}(\Theta) & \Sigma_{yx}(\Theta) \\ \Sigma_{xy}(\Theta) & \Sigma_{xx}(\Theta) \end{array} \right] \\ &= \left[\begin{array}{l} \Lambda_y(I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)[(I-B)^{-1}]'\Lambda_y^{'} + \Theta_\epsilon \\ \Lambda_x\Phi\Gamma'[(I-B)^{-1})]'\Lambda_y^{'} \end{array} \right. \end{aligned} \tag{19}$$

$$\begin{array}{l} \Lambda_y(I-B)^{-1}\Gamma\Phi\Lambda_x^{'} \\ \Lambda_x\Phi\Lambda_x^{'} + \Theta_\delta \end{array}$$

Die Identifikation der Modellparameter

Die Identifikation der Parameter im allgemeinen Strukturgleichungsmodell baut auf dem im Rahmen der Meßmodelle und der konfirmatorischen Faktorenanalyse erörterten Regelsystem auf. Zwei notwendige, aber nicht hinreichende, Bedingungen der Modellidentifikation sind für die Parameterschätzung zu erfüllen:

1. Die Anzahl der zu schätzenden Parameter im Parametervektor Θ muß gleich oder kleiner sein als die Anzahl der empirischen Parameter in der Kovarianzmatrix S.
2. Jede latente Variable (ξ und η) muß eine Skala erhalten.

Die erste Bedingung kann durch die t-Regel ($t \leq (p+q)(p+q+1)/2$) mit p als die Anzahl der x-Variablen und q als die Anzahl der y-Variablen erfüllt werden. Die zweite Bedingung wird entweder dadurch erfüllt, daß die Varianzen jeder latenten Variablen auf eine Konstante festgesetzt werden ($\phi_{11} = 1.0$ und $\psi_{11} = 1.0$) oder die latenten Variablen die Skalierung eines ihrer Indikatoren übernehmen ($\lambda_{x_{11}} = 1.0$ und $\lambda_{y_{11}} = 1.0$).

Regeln und Voraussetzungen zur Identifikation der Parameter in allgemeinen Strukturgleichungsmodellen

Identifikationsregel	Voraussetzungen	Charakterisierung
t-Regel	$t \leq \nu(\nu + 1)/2$	notwendig
Zwei-Schritt-Regel	1. Prüfung des Meßmodells ohne Strukturmodell 2. Prüfung des Strukturmodells ohne Meßmodell	hinreichend
MIMIC-Regel	Spezifikation gemäß Gleichung 5 Anzahl der x-Variablen ≥ 1 Anzahl der y-Variablen ≥ 2	hinreichend

t = Anzahl der freien Parameter

ν = Anzahl der gemessenen Variablen ($p + q$)

Standardisierte und unstandardisierte Koeffizienten

Da keine festen Intervallgrenzen für unstandardisierte Parameter existieren, kann bei einem Vergleich von Parametern im Modell nicht beurteilt werden, welcher Effekt am stärksten und welcher am schwächsten ist. Dieser Vergleich läßt sich mit standardisierten Koeffizienten durchführen, da diese eine feste Unter- und Obergrenze haben (zwischen -1 und +1).

Die standardisierten Koeffizienten β_{ij}^{s} und γ_{ij}^{s} des Strukturmodells werden wie folgt berechnet:

$$\beta_{ij}^{s} = \beta_{ij} \left(\frac{\sigma_{jj}}{\sigma_{ii}} \right)^{1/2} \tag{20}$$

$$\gamma_{ij}^{s} = \gamma_{ij} \left(\frac{\sigma_{jj}}{\sigma_{ii}} \right)^{1/2} \tag{21}$$

Das Superscript s kennzeichnet den standardisierten Koeffizienten, i bezieht sich auf die abhängige, j auf die unabhängige latente Variable. σ_{ii} und σ_{jj} sind die jeweiligen geschätzten Varianzen der abhängigen und unabhängigen latenten Variablen.

Die standardisierten Koeffizienten λ_{ij}^{s} des Meßmodells werden folgendermaßen berechnet:

$$\lambda_{ij}^{s} = \lambda_{ij} \left(\frac{\sigma_{jj}}{\sigma_{ii}} \right)^{1/2} \tag{22}$$

Hier bezieht sich i auf die manifeste Variable x oder y im Meßmodell, j auf die latente Variable ξ oder η. σ_{ii} ist dann die modellimplizierte Varianz der manifesten Variablen und σ_{jj} die geschätzte Varianz der jeweiligen latenten Variablen.

Gleichungen 20 bis 22 sind auf einzelne Parameter bezogen, in der Praxis werden aber ganze Parametermatrizen standardisiert.

Die Effektzerlegung

Tabelle 5 zeigt, daß fünf verschiedene Effektzerlegungen möglich sind, wobei sich diese auf zwei reduzieren können, wenn ein Strukturgleichungsmodell auf der latenten Ebene nur mit η-Variablen und auf der manifesten Ebene nur mit y-Variablen definiert wird.

Übersicht über die Effektzerlegungen im allgemeinen Strukturgleichungsmodell

Nr.	Kausalrichtung	Matrizen	Interpretation
1	$\xi \rightarrow \eta$	$(I - B)^{-1}\Gamma$	Totaler Kausaler Effekt von ξ nach η
2	$\eta \rightarrow \eta$	$(I - B)^{-1}$ - I	Totaler Kausaler Effekt von η nach η
3	$\xi \rightarrow x$	Λ_x	Totaler Kausaler Effekt von ξ nach x
4	$\xi \rightarrow y$	$\Lambda_y(I - B)^{-1}\Gamma$	Totaler Kausaler Effekt von ξ nach y
5	$\eta \rightarrow y$	$\Lambda_y(I - B)^{-1}$	Totaler Kausaler Effekt von η nach y

Die formale Grundlage für die Effektzerlegung bildet Gleichung 19.

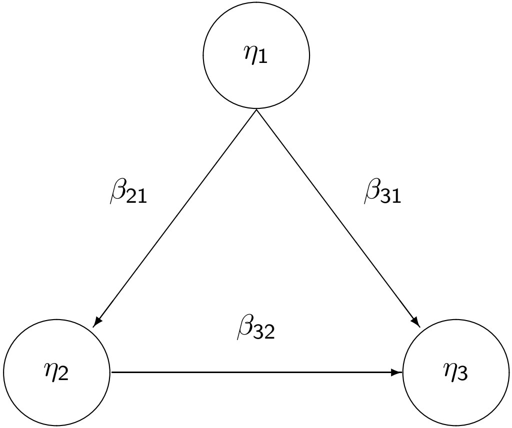
Im folgenden wird die Zerlegung der Koffizienten auf der latenten Ebene des Strukturgleichungsmodells betrachtet, da diese für inhaltliche Interpretationen am wichtigsten sind. Zur Vereinfachung werden alle latenten Variablen als η-Variablen definiert. Die totalen Effekte der latenten Variablen η lassen sich als infinite Summe einer geometrischen Reihe darstellen:

$$T_{\eta\eta} = \sum_{k=1}^{\infty} B^k \tag{23}$$

wobei die Matrix der totalen Effekte $T_{\eta\eta}$ nur dann definiert ist, wenn die Summe zu einer Matrix mit finiten Elementen konvergiert. Mit k werden die Effekte der Matrix B differenziert.

1. Bei $k = 1$ enthält Matrix B nur direkte Effekte.
2. Bei $k = 2$ enthält Matrix B die indirekten Effekte vermittelt über *eine* latente Variable η.
3. Bei $k = 3$ enthält Matrix B die indirekten Effekte vermittelt über *zwei* latente Variablen.
4. $\rightarrow$ bei $k > 1$ enthält Matrix B die jeweiligen indirekten Effekte der Länge k.

Strukturgleichungsmodell mit drei latenten Variablen (rekursiv)



Wenn ein rekursives Modell mit drei η-Variablen spezifiziert wird, dann beinhaltet Matrix B nach der Abbildung drei Parameter:

$$B = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \tag{24}$$

Setzt man diese Modellspezifikation in Gleichung 23 ein, dann können mit den ersten beiden Termen der infiniten Serie die totalen Effekte ermittelt werden:

$$\begin{aligned} T_{\eta\eta} &= B^1 + B^2 + B^3 \\ &= \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{21}\beta_{32} & 0 & 0 \end{pmatrix} + 0 \end{aligned} \tag{25}$$

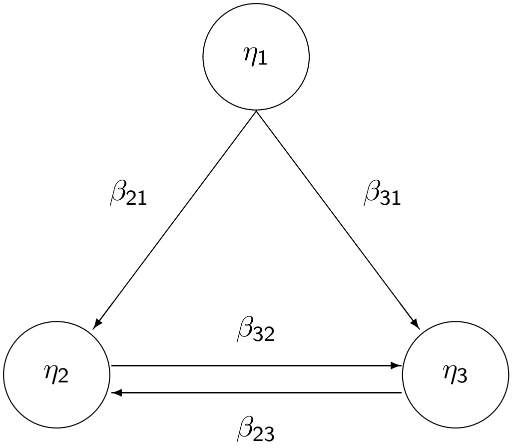
Es wird ersichtlich, daß für $k \geq 3$ alle B^k Null werden, die Serie konvergiert und somit die totalen kausalen Effekte bestimmt werden können.

Der erste Term B^1 enthält die drei direkten kausalen Effekte zwischen den drei latenten Variablen η_1, η_2 und η_3. Der zweite Term enthält den indirekten Effekt der Variablen η_1 auf η_3 vermittelt über η_2. Da keine indirekten Effekte über mehr als eine Variable existieren, ist $B^3 = 0$. Werden B^1 und B^2 aufsummiert, dann wird ersichtlich, daß der totale kausale Effekt immer die Summe der direkten und indirekten Effekte ist:

$$T_{\eta\eta} = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} + \beta_{21} & \beta_{32}\beta_{32} & 0 \end{pmatrix} \tag{26}$$

Für ein nicht-rekursives Modell ist die Situation etwas komplizierter. Es werden wieder drei η-Variablen spezifiziert, wobei η_2 auf η_3 wirkt als auch umgekehrt η_3 auf η_2.

Strukturgleichungsmodell mit drei latenten Variablen (nicht-rekursiv)



Matrix B beinhaltet nun vier Parameter:

$$B = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & \beta_{23} \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \tag{27}$$

$$\begin{aligned} T_{\eta\eta} &= B^1 + B^2 \\ &+ B^3 + B^4 \\ &= \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & \beta_{23} \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ \beta_{31}\beta_{23} & \beta_{32}\beta_{23} & 0 \\ \beta_{21}\beta_{32} & 0 & \beta_{32}\beta_{23} \end{pmatrix} \\ &+ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \beta_{32}\beta_{23}^2 \\ 0 & \beta_{32}^2\beta_{23} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \beta_{32}^2\beta_{23}^2 & 0 \\ 0 & 0 & \beta_{32}^2\beta_{23}^2 \end{pmatrix} \end{aligned}$$

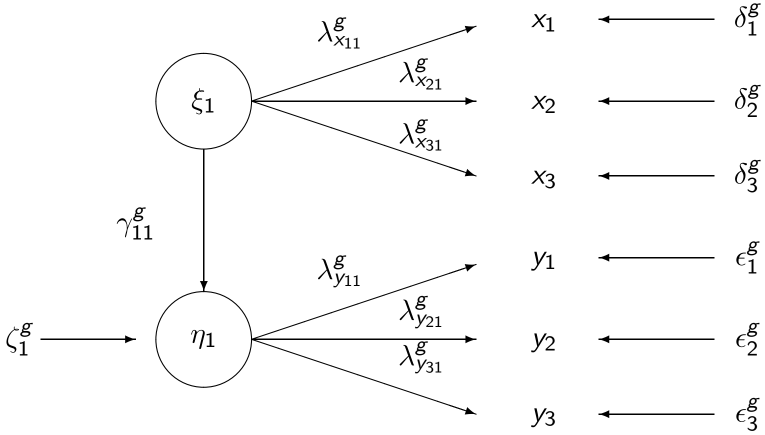
Im Unterschied zum rekursiven Modell kann im nicht-rekursiven Modell eine abhängige Variable einen Effekt auf sich selber haben. Die Diagonale in den Termen B^2, B^3 und B^4 enthält damit nicht nur Nullen. Daraus folgt, daß B^k nicht notwendigerweise Null ist, wenn k größer ist als die Anzahl der Variablen im Modell. Der Zyklus der indirekten Effekte einer Variablen auf sich selber kann in der geometrischen Reihe beliebig fortgesetzt werden. Die Konvergenz ist gewährleistet, wenn die Werte in B^k mit zunehmenden k gegen Null streben.
Werden alle Terme der Gleichung 5 aufsummiert, dann erhält man die totalen kausalen Effekte des Modells:

$$T_{\eta\eta} = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} + \beta_{31}\beta_{23} & \beta_{32}\beta_{23} + \beta_{32}^2\beta_{23}^2 & \beta_{23} + \beta_{32}\beta_{23}^2 \\ \beta_{31} + \beta_{21}\beta_{32} & \beta_{32} + \beta_{32}^2\beta_{23} & \beta_{32}\beta_{23} + \beta_{32}^2\beta_{23}^2 \end{pmatrix} \tag{28}$$

Multiple Gruppenvergleiche

Die folgende Abbildung zeigt ein Strukturgleichungsmodell für $(g = 1, 2 \ldots G)$-Gruppen:

Strukturgleichungsmodell mit zwei latenten Variablen für den multiplen Gruppenvergleich



Analog zu den Gleichungen 8, 9 und 10 läßt sich das Modell formal spezifizieren:

$$\eta_1 = \gamma_{11}^{g}\xi_1 + \zeta_1^{g} \tag{29}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda_{x_{11}}^{g} \\ \lambda_{x_{21}}^{g} \\ \lambda_{x_{31}}^{g} \end{pmatrix} * \begin{pmatrix} \xi_1 \end{pmatrix} + \begin{pmatrix} \delta_1^{g} \\ \delta_2^{g} \\ \delta_3^{g} \end{pmatrix} \tag{30}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \lambda_{y_{11}}^{g} \\ \lambda_{y_{21}}^{g} \\ \lambda_{y_{31}}^{g} \end{pmatrix} * \begin{pmatrix} \eta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1^{g} \\ \epsilon_2^{g} \\ \epsilon_3^{g} \end{pmatrix} \tag{31}$$

Es gelten auch hier die unter Gleichung 10 genannten Annahmen über die Kovariationen zwischen Meßfehlern (δ, ϵ) und latenten Variablen und die Erwartungswerte der Meßfehler.
Der multiple Gruppenvergleich des Strukturgleichungsmodells läßt auch einen simultanen Vergleich der Mittelwerte zu. Hierzu werden die Mittelwertsvektoren der Variablen für die einzelnen Gruppen berechnet und als empirische Größen dem Gruppenvergleichsmodell hinzugefügt.

Gleichungen 29, 30 und 31 werden entsprechend um die Mittelwertsvektoren α, τ_x und τ_y erweitert:

$$\eta_1 = \alpha + \gamma_{11}^g \xi_1 + \zeta_1^g \tag{32}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \tau_{x_1} \\ \tau_{x_2} \\ \tau_{x_3} \end{pmatrix} + \begin{pmatrix} \lambda_{x_{11}}^g \\ \lambda_{x_{21}}^g \\ \lambda_{x_{31}}^g \end{pmatrix} * \begin{pmatrix} \xi_1 \end{pmatrix} + \begin{pmatrix} \delta_1^g \\ \delta_2^g \\ \delta_3^g \end{pmatrix} \tag{33}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \tau_{y_1} \\ \tau_{y_2} \\ \tau_{y_3} \end{pmatrix} + \begin{pmatrix} \lambda_{y_{11}}^g \\ \lambda_{y_{21}}^g \\ \lambda_{y_{31}}^g \end{pmatrix} * \begin{pmatrix} \eta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1^g \\ \epsilon_2^g \\ \epsilon_3^g \end{pmatrix} \tag{34}$$

Dieses Strukturgleichungsmodell ist für jede Gruppe g durch die Parameter in den Gleichungen definiert, wobei mit $g = 1, 2, \ldots G$ die jeweilige Gruppe bezeichnet wird. Es gelten auch hier die unter Gleichung 10 genannten Annahmen. Mit τ_x werden die Mittelwerte der manifesten Variablen x bezeichnet, mit τ_y die Mittelwerte der manifesten Variablen y. Der Mittelwertsvektor der latenten Variablen ξ wird mit κ bezeichnet, der entsprechende Vektor der latenten Variablen η mit α.

Da die manifesten Mittelwertsinformationen im Gruppenvergleich berücksichtigt werden, läßt sich auch auf der latenten Ebene eine Mittelwertsschätzung vornehmen. Allerdings sind nur die Mittelwertsdifferenzen zwischen den Gruppen schätzbar. Hierzu werden in einer Gruppe die latenten Mittelwerte auf Null restringiert:

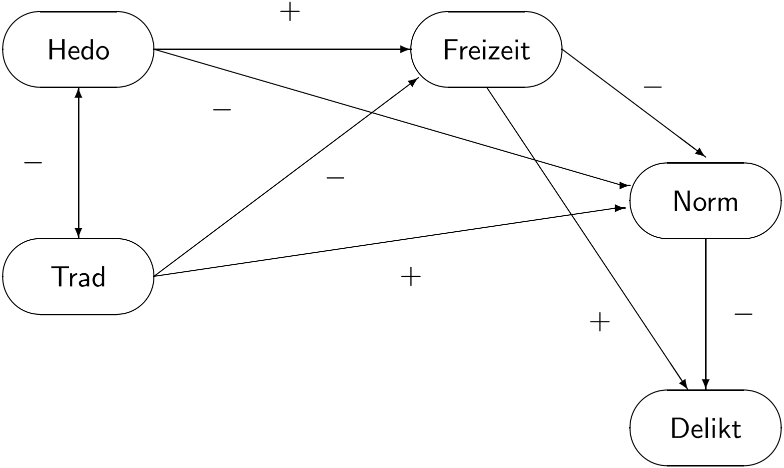
$$\begin{aligned} \kappa_1^1 &= 0 \\ \alpha_1^1 &= 0 \end{aligned} \tag{35}$$

Demnach sind κ_1^2 und α_1^2 zwei zu schätzende Parameter, die als Differenzen zu Null zu interpretieren sind. Die Mittelwerte der manifesten Variablen sind über die Gruppen gleichzusetzen:

$$\begin{aligned} \tau_{x_1}^1 &= \tau_{x_1}^2 \\ \tau_{x_2}^1 &= \tau_{x_2}^2 \\ \tau_{x_3}^1 &= \tau_{x_3}^2 \\ \tau_{y_1}^1 &= \tau_{y_1}^2 \\ \tau_{y_2}^1 &= \tau_{y_2}^2 \\ \tau_{y_3}^1 &= \tau_{y_3}^2 \end{aligned} \tag{36}$$

Empirische Beispiele

Strukturgleichungsmodell zum Zusammenhang von Wertorientierungen, Freizeit, Rechtsnormen und Delinquenz (Richtung der Hypothesen)



Folgende Hypothesen werden dem Strukturgleichungsmodell zu Grunde gelegt:

1. Je stärker eine hedonistische Werteinstellung geäußert wird, desto eher neigen die Jugendlichen zu Freizeitverhalten, das außerhalb des familiären Umfeldes stattfindet.
2. Je stärker die hedonistische Werteinstellung vorliegt, desto weniger akzeptieren Jugendliche Rechtsnormen.
3. Je stärker eine traditionelle Werteinstellung geäußert wird, desto weniger neigen die Jugendlichen zu Freizeitverhalten, das außerhalb des familiären Umfeldes stattfindet.
4. Je stärker die traditionelle Werteinstellung vorliegt, desto eher akzeptieren Jugendliche Rechtsnormen.
5. Je stärker die nach außen gerichteten Freizeitorientierungen sind, desto eher werden Delikte von Jugendlichen begangen.
6. Je stärker die Freizeitorientierungen, desto weniger Akzeptanz von Rechtsnormen.
7. Je weniger Akzeptanz von Rechtnormen, desto eher werden Delikte von Jugendlichen begangen.

Es wird keine Richtung für den Zusammenhang zwischen den beiden Wertorientierungen (*Hedo* und *Trad*) angenommen. Beide Wertorientierungen sollten negativ miteinander korrelieren, da anzunehmen ist, daß mit zunehmenden hedonistische Orientierungen die traditionellen Werte abnehmen und umgekehrt. Eine direkte Wirkung der Wertorientierungen auf die Deliktraten wird nicht postuliert. Damit wird angenommen, daß der gesamte Einfluß der Wertorientierungen über das Freizeitverhalten und die Akzeptanz von Rechtsnormen vermittelt wird. Alle aufgestellten Hypothesen beziehen sich auf die latente Ebene des Strukturgleichungsmodells und werden im folgenden empirisch überprüft. Der Modelltest wird anhand der Daten des 8. Schuljahrgangs aller weiterführenden Schulen der Stadt Münster aus dem Jahre 2001 vorgenommen (N = 1942).

- Der Faktor der hedonistischen Werteinstellung (*Hedo*) wird durch je drei Items gemessen:
    1. Das erste Items der hedonistischen Werteinstellung drückt die Meinung aus, daß der Sinn des Lebens darin besteht, Spaß zu haben und kaufen zu können, was einem gefällt (I0053).
    2. Das zweite Item spiegelt die Meinung wieder, man hätte großes Verständnis für Leute, die nur täten, wozu sie Lust hätten (I0054).
    3. Das dritte Item verdeutlicht die Devise: Genießen und möglichst angenehm leben (I0076).
- Der Faktor der traditionellen Werteinstellung (*Trad*) wird durch je drei Items gemessen:
    1. Nur durch Pflichterfüllung könne man das Lebensziel erreichen (I0061).
    2. In der Schule solle man sich nichts zu Schulden kommen lassen (I0066).
    3. Generell haben alte Werte eine große Bedeutung (I0083).

- Zwei Messungen stehen für die nach außen gerichteten Freizeitorientierungen zur Verfügung:
  1. Auf Partys gehen (l0618) und
  2. Abhängen (l0619).
- Für die Akzeptanz von Rechtsnormen sind drei Items ausgewählt worden:
  1. Es ist wichtig, die Gesetze zu beachten (t0722).
  2. Man schadet anderen, die nichts dafür können (t0725).
  3. Ich schade mir selbst dabei (t0727).
- Die Prävalenz- und Inzidenzraten aller für die vorangegangenen zwölf Monate berichteten Delikte (im folgenden abgekürzt *Prae* und *Inz*) bilden die Messungen für das zentrale abhängige Konstrukt, die Dunkelfelddelikte der befragten Jugendlichen (*Delikt*).

Die Momente der Verteilungen für die Messungen der Konstrukte

Konstrukt	Item/Index	$\bar{x}$	s	s^3	s^4
Hedo	I0053	2.753	0.934	-0.116	-0.981
	I0054	2.508	0.951	0.072	-0.923
	I0076	2.771	0.873	-0.215	-0.692
Trad	I0061	3.019	0.840	-0.533	-0.352
	I0066	2.768	0.840	-0.276	-0.492
	I0083	2.580	0.903	-0.062	-0.780
Freizeit	I0618	2.558	0.926	0.160	-0.900
	I0619	2.604	1.040	-0.120	-1.157
Norm	t0722	2.864	0.939	-0.503	-0.606
	t0725	3.002	0.963	-0.694	-0.483
	t0727	3.023	0.974	-0.705	-0.526
Delikt	Prae	1.308	2.470	3.043	11.480
	Inz	1.188	1.674	0.889	-1.028

$\bar{x}$ = Mittelwert; s = Standardabweichung; s^3 = Schiefe; s^4 = Kurtosis

Korrelationsmatrix für die Messungen der Werteinstellungen, des Freizeitverhaltens, der Norm und der Delinquenz

	Hedo			Trad			Freizeit	
	l0053	l0054	l0076	l0061	l0066	l0083	l0618	l0619
l0053	1.000							
l0054	**0.461**	1.000						
l0076	**0.460**	**0.312**	1.000					
l0061	0.018	-0.048	0.000	1.000				
l0066	-0.014	-0.040	0.055	**0.295**	1.000			
l0083	-0.036	-0.076	0.064	**0.297**	**0.339**	1.000		
l0618	0.295	0.220	0.210	0.006	-0.069	-0.017	1.000	
l0619	0.272	0.221	0.207	-0.044	-0.075	-0.084	**0.455**	1.000
t0722	-0.184	-0.220	-0.184	0.210	0.189	0.223	-0.218	-0.248
t0725	-0.197	-0.207	-0.128	0.124	0.107	0.123	-0.153	-0.143
t0727	-0.162	-0.167	-0.122	0.164	0.185	0.151	-0.184	-0.144
Prae	0.251	0.254	0.225	-0.093	-0.147	-0.148	0.365	0.324
Inz	0.283	0.247	0.228	-0.055	-0.133	-0.122	0.369	0.357

	Norm			Delikt	
	t0722	t0725	t0727	Prae	Inz
t0722	1.000				
t0725	**0.465**	1.000			
t0727	**0.462**	0.542	1.000		
Prae	-0.354	-0.282	-0.322	1.000	
Inz	-0.356	-0.264	-0.308	**0.728**	1.000

Das Meßmodell wird für die unabhängigen latenten Variablen (*Hedo* und *Trad*) entsprechend Gleichung 2 spezifiziert:

$$\begin{pmatrix} I0053 \\ I0054 \\ I0076 \\ I0061 \\ I0066 \\ I0083 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{pmatrix} * \begin{pmatrix} Hedo \\ Trad \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix} \qquad (37)$$

Um die Identifikation des Meßmodells zu gewährleisten bzw. die latenten Variablen zu skalieren, werden die Parameter λ_{11} und λ_{42} auf den Wert 1.0 fixiert.

Die Meßfehlervarianzen werden in der Diagonalen der Matrix Θ_δ angegeben. Nach Gleichung 3 wird das Meßmodell für die abhängigen latenten Variablen (*Freizeit*, *Norm* und *Delikt*) spezifiziert:

$$\begin{pmatrix} l0618 \\ l0619 \\ t0722 \\ t0725 \\ t0727 \\ Prae \\ Inz \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ 0 & \lambda_{32} & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & 0 & \lambda_{63} \\ 0 & 0 & \lambda_{73} \end{pmatrix} * \begin{pmatrix} Freiz \\ Norm \\ Delikt \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \end{pmatrix} \quad (38)$$

Auch hier wird zur Identifikation je eine Faktorenladung pro latente Variable auf den Wert 1.0 festgesetzt (λ_{11}, λ_{32} und λ_{63}). Die Meßfehlervarianzen sind in der Diagonalen der Matrix Θ_ϵ enthalten. Zwischen Item $t0725$ und $t0727$ wird eine Meßfehlerkovariation angenommen ($\theta_{\delta_{54}}$), da in beiden Itemkonnotationen das Wort „schaden“ vorkommt und hier systematische Antworteffekte zu erwarten sind. Nach Gleichung 1 werden nun die Beziehungen der latenten Variablen im Strukturmodell spezifiziert:

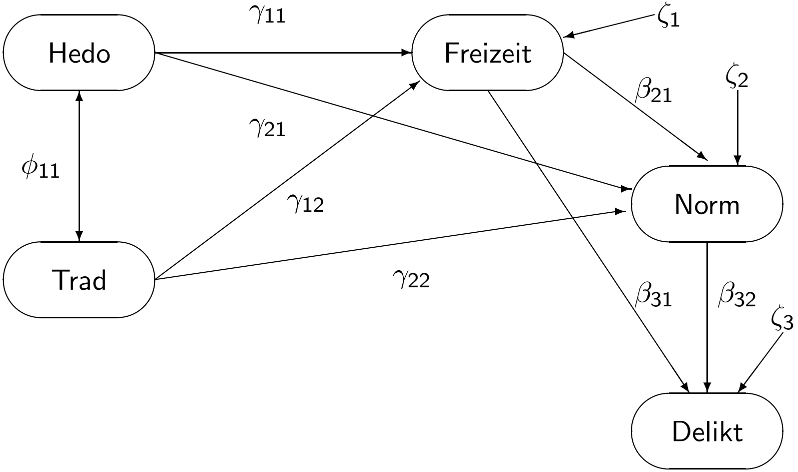
$$\begin{pmatrix} \mathit{Freiz} \\ \mathit{Norm} \\ \mathit{Delikt} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} * \begin{pmatrix} \mathit{Freiz} \\ \mathit{Norm} \\ \mathit{Delikt} \end{pmatrix} \tag{39}$$

$$+ \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ 0 & 0 \end{pmatrix} * \begin{pmatrix} \mathit{Hedo} \\ \mathit{Trad} \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix}$$

In der Matrix B sind drei Parameter spezifiziert, in der Matrix Γ vier Parameter. Direkte Beziehungen zwischen den Wertorientierungen und den Deliktḧaufigkeiten der Befragten werden explizit ausgeschlossen (vgl. die letzte Zeile der Matrix Γ in Gleichung 39). Die Varianzen und Kovarianzen der unabhängigen latenten Variablen werden in der Matrix Φ spezifiziert, die Residualvarianzen der latenten abhängigen Variablen in der Diagonalen der Matrix Ψ:

$$\mathbf{\Phi} = \begin{pmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{pmatrix} \qquad \mathbf{\Psi} = \begin{pmatrix} \psi_{11} & & \\ 0 & \psi_{22} & \\ 0 & 0 & \psi_{33} \end{pmatrix} \tag{40}$$

Spezifikation des Strukturmodells zum Zusammenhang von Wertorientierungen, Freizeit, Rechtsnormen und Delinquenz



In den beiden Meßmodellen sind 22 Parameter zu schätzen (je 4 Parameter aus den Matrizen Λ_x und Λ_y, 6 Parameter aus der Matrix Θ_δ und 8 Parameter aus der Matrix Θ_ϵ).
Im Strukturmodell sind 13 Parameter zu schätzen (je 3 Parameter aus den Matrizen B, Φ und Ψ und 4 Parameter aus der Matrix Γ).
Die Freiheitsgrade betragen $df = 91 - 35 = 56$. Um die höheren Momente für die Korrektur der inferenzstatistischen Größen zu berücksichtigen, sind die Modelle mit der üblichen ML-Diskrepanzfunktion und der asymptotischen Varianz-/Kovarianzmatrix berechnet worden. Die empirische Kovarianzmatrix S basiert auf $N = 1549$ Personen. Aus Vergleichsgründen werden auch die Parameter der WLS-Diskrepanzfunktion aufgeführt.

Modellanpassungen des Strukturgleichungsmodells (ML- und WLS-Schätzfunktion)

Modell (ML)	χ^2	df	RMSEA	p-Value	GFI
C1	113.08	56	0.026	1.00	0.989
C2	113.48				
C3	98.38		0.022	1.00	0.989
C4	106.17				
Modell (WLS)	χ^2	**df**	**RMSEA**	**p-Value**	**GFI**
C1	106.50	56	0.024	0.989	0.992

C1: $(N-1)\cdot F_D$ mit D als Diskrepanzfunktion ML oder WLS.

C2: $(N-1)\cdot F_D$ mit D als Diskrepanzfunktion WLS, wobei die Gewichtungsmatrix W unter der Multinormalverteilungsbedingung benutzt wird.

C3: *Satorra-Bentler scaled chi-square statistic* (Satorra & Bentler, 1988).

C4: *Satorra chi-square statistic* (Satorra, 1993), wobei die asymptotische Varianz-/Kovarianzmatrix als Gewichtungsmatrix W benutzt wird.

Unstandardisierte Faktorenladungen (λ), Meßfehler (δ) und die jeweiligen Standardfehler (z-Wert) der Meßmodelle

Meßmodell: Hedo und Trad						
ML-Schätzung						
Item	λ	z-Wert	z-Wert(r)	δ	z-Wert	z-Wert(r)
I0053	1.00			0.34	21.43	11.06
I0054	0.77	17.08	17.66	0.59	22.04	22.81
I0076	0.69	16.87	16.19	0.51	22.48	19.63
I0061	1.00			0.52	20.76	19.46
I0066	1.09	11.58	11.01	0.48	18.61	17.39
I0083	1.20	11.59	11.12	0.54	17.96	16.87
WLS-Schätzung						
Item	λ	z-Wert		δ	z-Wert	
I0053	1.00			0.30	9.81	
I0054	0.76	18.32		0.56	22.44	
I0076	0.66	16.86		0.51	20.87	
I0061	1.00			0.51	19.24	
I0066	1.06	10.77		0.46	17.47	
I0083	1.14	10.97		0.55	18.29	

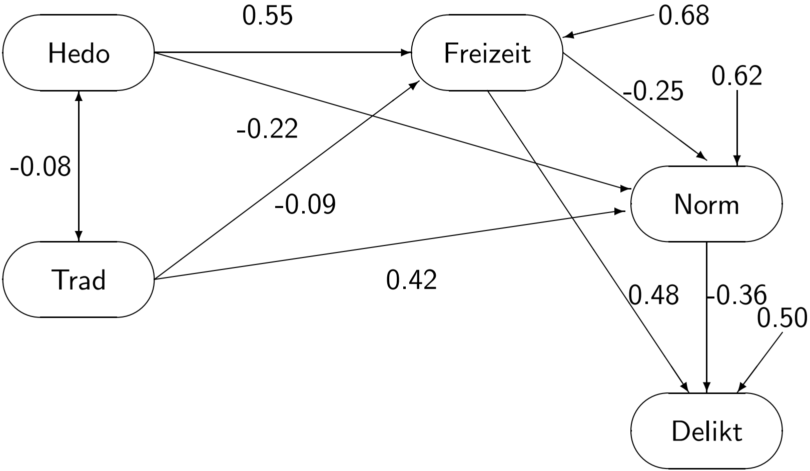
Unstandardisierte Faktorenladungen (λ), Meßfehler (ϵ) und die jeweiligen Standardfehler (z-Wert) der Meßmodelle

Meßmodell: Freizeit, Norm und Delikt						
ML-Schätzung						
Item	λ	z-Wert	z-Wert(r)	ϵ	z-Wert	z-Wert(r)
l0618	1.00			0.45	16.38	16.18
l0619	1.06	16.46	16.57	0.62	18.48	17.90
t0722	1.00			0.37	12.82	11.12
t0725	0.80	16.10	15.21	0.60	20.89	17.34
t0727	0.83	16.52	15.72	0.59	20.19	17.88
Prae	1.00			1.74	12.02	6.82
Inz	0.69	27.92	20.39	0.72	10.77	10.28
WLS-Schätzung						
Item	λ	z-Wert		ϵ	z-Wert	
l0618	1.00			0.44	16.40	
l0619	1.08	17.49		0.58	17.05	
t0722	1.00			0.34	10.70	
t0725	0.79	15.16		0.58	17.26	
t0727	0.85	15.98		0.55	17.07	
Prae	1.00			1.47	6.46	
Inz	0.73	21.64		0.67	10.01	

Unstandardisierte Regressionskoeffizienten (β, γ) und die jeweiligen Standardfehler (z-Wert) des Strukturmodells

Strukturmodell (ML)						
Konstrukt	Koeff.	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit	β,γ	0.481	-0.132			
	z-Wert	12.860	-2.299			
	z-Wert(r)	13.079	-2.243			
Norm	β,γ	-0.215	0.682	-0.278		
	z-Wert	-4.640	9.103	-4.915		
	z-Wert(r)	-4.012	8.757	-4.362		
Delikt	β,γ			1.555	-1.044	
	z-Wert			11.560	-9.694	
	z-Wert(r)			10.583	-7.775	
Strukturmodell (WLS)						
Konstrukt	Koeff.	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit	β,γ	0.468	-0.093			
	z-Wert	13.680	-1.595			
Norm	β,γ	-0.167	0.664	-0.308		
	z-Wert	-3.400	8.665	-5.127		
Delikt	β,γ			1.511	-0.934	
	z-Wert			11.284	-7.956	

Standardisierte Lösung des Strukturmodells (ML-Parameter)



Standardisierte indirekte und totale Effekte des Strukturmodells (ML-Parameter)

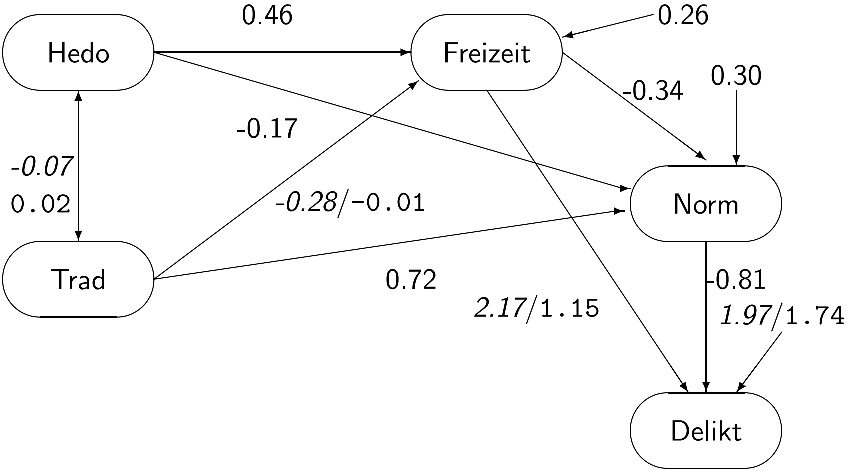
Indirekte Effekte					
Konstrukt	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit					
Norm	-0.14	0.02			
Delikt	0.39	-0.20	0.09		
Totale Effekte					
Konstrukt	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit	0.55	-0.09			
Norm	-0.35	0.44	-0.25		
Delikt	0.39	-0.20	0.56	-0.36	

Indirekte und die in Abbildung 7 eingetragenen direkten Effekte summieren sich zu den jeweiligen totalen Effekten.

Vergleich der Modellvarianten nach dem multiplen Gruppenvergleich

Modell	Gruppe		χ^2	df	χ^2_{Diff}	df_{Diff}	RMSEA	GFI
Variante 1	m		221.77					0.954
Basismodell	w		284.53					0.950
		$\sum$	506.31	147	$--$	$--$	0.056	
Variante 2	m		157.63					0.967
Θ_δ,Θ_ϵ	w		133.54					0.975
		$\sum$	291.17	133	215.14	14	0.040	
Variante 3	m		117.27					0.976
B, Γ,Φ,Ψ	w		108.76					0.979
		$\sum$	226.03	129	65.14	4	0.032	
Variante 4	m		101.42					0.979
Λ_x,Λ_y	w		91.19					0.983
		$\sum$	192.61	121	33.42	8	0.028	

Ergebnis der Modellvariante 3 des Strukturmodells



xx = Parameter der ersten Gruppe (männlich)

`xx` = Parameter der zweiten Gruppe (weiblich)

Mittelwerte der manifesten Variablen (τ_x und τ_y) und Mittelwertsdifferenzen der latenten Variablen (κ und α) für die Modellvariante 3

Modellvariante 3					
Item	τ_x	z-Wert	Item	τ_y	z-Wert
I0053	2.82	86.31	I0618	2.46	77.69
I0054	2.55	85.91	I0619	2.50	70.31
I0076	2.81	102.91	t0722	2.75	81.14
I0061	3.04	114.03	t0725	2.93	94.37
I0066	2.80	101.90	t0727	2.94	90.59
I0083	2.62	89.53	Prae	1.72	17.95
			Inz	1.49	22.00
Faktor	κ	z-Wert	Faktor	α	z-Wert
Hedo	-0.13	2.98	Freizeit	0.24	6.17
Trad	-0.06	-1.84	Norm	0.30	6.87
			Delikt	-0.87	-8.49

Kategoriales Meßniveau in Strukturgleichungsmodellen

Wenn die kategorialen, manifesten Variablen im Strukturgleichungsmodell ordinales Meßniveau aufweisen, dann können polychorische bzw. polyserielle Korrelationskoeffizienten als Datenbasis dienen. Eine polyserielle Korrelation wird dann berechnet, wenn eine manifeste Variable kategoriales und die andere metrisches Meßniveau hat. Zentral ist die Annahme, daß die manifesten Variablen Messungen von latenten metrischen Indikatorvariablen sind, für die eine Normalverteilung unterstellt wird. Die Beziehung zwischen zwei metrischen Indikatorvariablen x^* und y^* wird über ein Schwellenwertmodell beschrieben, mit dem die polychorische bzw. die polyserielle Korrelation geschätzt werden kann. Das Schwellenwertmodell beschreibt die Beziehung zwischen den Indikatorvariablen (x^* bzw. y^*) und den ordinalen Variablen (x und y).

Das in den Gleichungen 2 und 3 für metrische manifeste Variablen formulierte Meßmodell wird an das ordinale Meßmodell angepasst:

$$x^* = \Lambda_x \xi + \delta \tag{41}$$

$$y^* = \Lambda_y \eta + \epsilon \tag{42}$$

Σ^* ist die geschätzte Varianz-/Kovarianzmatrix der Indikatorvariablen x^* und y^*, wobei gilt: $\Sigma^* = \Sigma(\Theta)$. Gegenüber Strukturgleichungsmodellen mit metrischen, manifesten Variablen wird also nur die Matrix Σ durch Σ^* ersetzt.
Polychorische und polyserielle Korrelationen können aus den Rohdaten mit dem Programm PRELIS (Jöreskog & Sörbom, 1993c) berechnet werden. Parallel dazu können die höheren Momente durch die asymptotische Varianz-/Kovarianzmatrix (ebenfalls mit PRELIS) ermittelt werden, so daß eine Schätzung der Parameter des Strukturgleichungsmodells auf der Basis der polychorischen, polyseriellen und Produkt-Moment-Korrelationen mit der WLS-Diskrepanzfunktion anzustreben ist.

Mittelwerte, Standardabweichungen und Schwellenwerte der Indikatorvariablen

Variable	$\bar{x}$	s	τ			
I0618	0.000	1.000	-1.231	0.074	0.855	
I0619	0.000	1.000	-0.912	-0.113	0.706	
I0053	0.000	1.000	-1.375	-0.201	0.649	
I0054	0.000	1.000	-1.037	0.048	0.927	
I0076	0.000	1.000	-1.440	-0.327	0.778	
I0061	0.000	1.000	-1.654	-0.689	0.485	
I0066	0.000	1.000	-1.459	-0.384	0.874	
I0083	0.000	1.000	-1.168	-0.092	0.976	
t0722	0.000	1.000	-1.242	-0.513	0.598	
t0725	0.000	1.000	-1.267	-0.651	0.353	
t0727	0.000	1.000	-1.274	-0.643	0.295	
Inz	0.000	1.000	0.293	0.489	0.604	0.780

	Hedo			Trad			Freizeit	
	l0053	l0054	l0076	l0061	l0066	l0083	l0618	l0619
l0053	1.000							
l0054	**0.524**	1.000						
l0076	**0.529**	**0.361**	1.000					
l0061	0.032	-0.052	0.009	1.000				
l0066	-0.012	-0.046	-0.058	**0.352**	1.000			
l0083	-0.043	-0.090	-0.075	**0.354**	**0.391**	1.000		
l0618	0.339	0.249	0.242	0.010	-0.076	-0.020	1.000	
l0619	0.317	0.258	0.240	-0.049	-0.088	-0.100	**.523**	1.000
t0722	-0.208	-0.252	-0.206	0.255	0.221	0.264	-0.249	-0.289
t0725	-0.231	-0.241	-0.150	0.159	0.127	0.156	-0.174	-0.176
t0727	-0.189	-0.196	-0.136	0.205	0.220	0.186	-0.203	-0.172
Prae	0.279	0.273	0.251	-0.093	-0.153	-0.161	0.389	0.364
Inz	0.367	0.317	0.294	-0.059	-0.166	-0.151	0.464	0.462

	Norm			Delikt	
	t0722	t0725	t0727	Prae	Inz
t0722	1.000				
t0725	**0.541**	1.000			
t0727	**0.537**	**0.632**	1.000		
Prae	-0.363	-0.288	-0.327	1.000	
Inz	-0.429	-0.335	-0.382	**0.772**	1.000

Unstandardisierte Faktorenladungen (λ), Meßfehler (δ) und die jeweiligen Standardfehler (z-Wert) der Meßmodelle (WLS-Parameter)

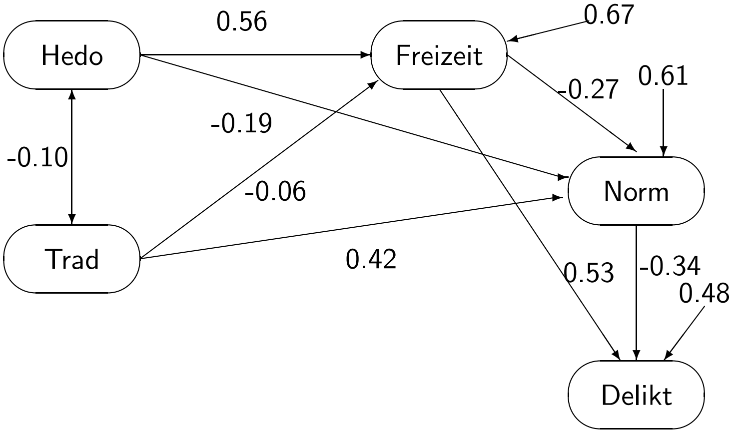
Meßmodell: Hedo und Trad				
Item	λ	z-Wert	δ	z-Wert
I0053	1.00		0.26	5.73
I0054	0.76	20.64	0.58	14.53
I0076	0.72	18.72	0.62	15.33
I0061	1.00		0.65	14.08
I0066	1.03	12.46	0.63	13.51
I0083	1.03	12.69	0.63	13.77

Meßmodell: Freizeit, Norm und Delikt				
Item	λ	z-Wert	ϵ	z-Wert
I0618	1.00		0.46	10.63
I0619	0.99	19.33	0.47	10.77
t0722	1.00		0.30	6.20
t0725	0.79	17.22	0.57	12.79
t0727	0.85	18.31	0.53	11.99
Prae	1.00		0.37	6.28
Inz	1.23	20.46	0.06	0.91

Unstandardisierte Regressionskoeffizienten (β, γ) und die jeweiligen Standardfehler (z-Wert) des Strukturmodells (WLS-Parameter)

Strukturmodell (WLS)						
Konstrukt	Koeff.	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit	β,γ	0.481	-0.078			
	z-Wert	14.706	-1.620			
Norm	β, γ	-0.186	0.595	-0.307		
	z-Wert	-3.690	9.605	-5.037		
Delikt	β, γ			0.576	-0.319	
	z-Wert			12.550	-7.702	

Standardisierte Lösung des Strukturmodells (WLS-Parameter)



Standardisierte indirekte und totale Effekte des Strukturmodells (WLS-Parameter)

Indirekte Effekte					
Konstrukt	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit					
Norm	-0.15	0.02			
Delikt	0.42	-0.18	0.09		

Totale Effekte					
Konstrukt	Hedo	Trad	Freizeit	Norm	Delikt
Freizeit	0.56	-0.06			
Norm	-0.34	0.44	-0.27		
Delikt	0.42	-0.18	0.63	-0.34	

Strukturgleichungsmodell als Markov-Modell

Die Anwendung von Strukturgleichungsmodellen zur Analyse von Paneldaten ist eng mit dem Ziel verbunden, Stabilität und Veränderung von theoretischen Konstrukten (latenten Variablen) und ihren Indikatoren (manifeste Variablen) zu untersuchen:

- Differenzierung nach *wahrer* Veränderung (*true change*) der latenten Variablen und nach unsystematischen Meßfehlern variierendem Antwortverhalten.
- Die Variabilität der Wiederholungsmessungen impliziert nicht unbedingt die zeitliche Konstanz bzw. Veränderung der latenten Variablen.

Werden Strukturgleichungsmodelle für Längsschnittdaten konstruiert, dann können die Stabilitäten zwischen den Variablen mit unterschiedlicher Restriktionen getestet werden. Drei grundlegende Situationen der Modellbildung sind zu unterscheiden:

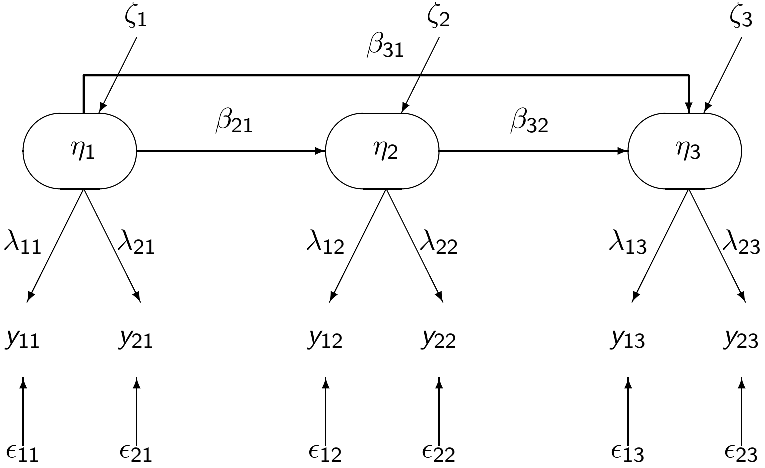
1. Perfekte Stabilität
2. Sokratische Veränderung
3. Markov-Struktur

Veränderung zwischen allen Meßzeitpunkten (Markov-Struktur)

$$
\begin{tikzcd}
X_{t_1} \arrow[r, "\beta_{t_2,t_1}"] \arrow[d, "p_{xX}"] & X_{t_2} \arrow[r, "\beta_{t_3,t_2}"] \arrow[d, "p_{xX}"] & X_{t_3} \arrow[d, "p_{xX}"] \\
x_{t_1} & x_{t_2} & x_{t_3} \\
\epsilon_x \arrow[u] & \epsilon_x \arrow[u] & \epsilon_x \arrow[u]
\end{tikzcd}
$$

Beim Markov-Modell wird eine Veränderung der latenten Variablen X zwischen den Meßzeitpunkten postuliert (β_{t_2,t_1}, β_{t_3,t_2}), wobei die aktuellen Zustandswahrscheinlichkeiten nur von denen zum unmittelbar vorgelagerten Zeitpunkt abhängig sind. Diese Modellierung wird auch als Markov-Modell erster Ordnung bezeichnet.

Latentes Markov-Modell zweiter Ordnung



$$\eta_t = B\eta_{t-1} + \zeta_t \qquad \text{mit} \qquad t = 1, 2, 3 \tag{43}$$

$$y_t = \Lambda_{y_t}\eta_t + \epsilon_t \qquad \text{mit} \qquad t = 1, 2, 3 \tag{44}$$

Restriktionen

Zur Identifikation:

$$\psi_{11} = \psi_{22} = \psi_{33} = 1.0$$

oder

$$\lambda_{y_{11}} = \lambda_{y_{12}} = \lambda_{y_{13}} = 1.0$$

Zeitstabilität der Faktorenladungen:

$$\lambda_{y_{11}} = \lambda_{y_{12}} = \lambda_{y_{13}}$$
$$\lambda_{y_{21}} = \lambda_{y_{22}} = \lambda_{y_{23}}$$

Zeitstabilität der Meßfehlervarianzen:

$$\sigma_{\epsilon_{11}} = \sigma_{\epsilon_{12}} = \sigma_{\epsilon_{13}}$$
$$\sigma_{\epsilon_{21}} = \sigma_{\epsilon_{22}} = \sigma_{\epsilon_{23}}$$

Autokovariationen der Meßfehler

$$\Theta_\epsilon = \begin{pmatrix} \sigma_{\epsilon_{11}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\epsilon_{21}} & 0 & 0 & 0 & 0 \\ \sigma_{\epsilon_{12,11}} & 0 & \sigma_{\epsilon_{12}} & 0 & 0 & 0 \\ 0 & \sigma_{\epsilon_{22,21}} & 0 & \sigma_{\epsilon_{22}} & 0 & 0 \\ \sigma_{\epsilon_{13,11}} & 0 & \sigma_{\epsilon_{13,12}} & 0 & \sigma_{\epsilon_{13}} & 0 \\ 0 & \sigma_{\epsilon_{23,21}} & 0 & \sigma_{\epsilon_{23,22}} & 0 & \sigma_{\epsilon_{23}} \end{pmatrix} \tag{45}$$

Zeitstabilität der Autokovariationen:

$$\sigma_{\epsilon_{12,11}} = \sigma_{\epsilon_{13,11}} = \sigma_{\epsilon_{13,12}}$$
$$\sigma_{\epsilon_{22,21}} = \sigma_{\epsilon_{23,21}} = \sigma_{\epsilon_{23,22}}$$

Beispiel

- Variablen und Daten für das Strukturgleichungsmodell im Längsschnitt sind der Ausländerstichprobe des Sozio-ökonomischen Panels (SOEP) entnommen. Die Stichprobe beinhaltet Ausländer, die der Gruppe der sogenannten Gastarbeiter zugeordnet werden können (Türken, Ex-Jugoslawen, Griechen, Italiener und Spanier).

- Zwei gemessene Variablen sollen die latente Dimension *Identität* konstruktvalide über drei Zeitpunkte erfassen:

  1. Die Variable x_{1t} $(t = 1, \ldots 3)$ mißt das *Zugehörigkeitsgefühl zu Deutschland*
     (SOEP: bp90a01, cp81a01, dp83a01).
  2. Die Variable x_{2t} $(t = 1, \ldots 3)$ mißt das *Zugehörigkeitsgefühl zur eigenen Nationalität*
     (SOEP: bp90a02, cp81a02, dp83a02).

Mittelwerte und Standardabweichungen der gemessenen Variablen aus dem SOEP für die Jahre 1985, 1986 und 1987

	$\bar{y}_{it}$	s_{it}
bp90a01 (y_{11})	3.963	1.117
bp90a02 (y_{21})	4.283	1.033
cp81a01 (y_{12})	3.964	1.098
cp81a02 (y_{22})	4.232	1.031
dp83a01 (y_{13})	3.953	1.089
dp83a02 (y_{23})	4.236	0.993

$N = 1844$ (listenweiser Ausschluß fehlender Werte)

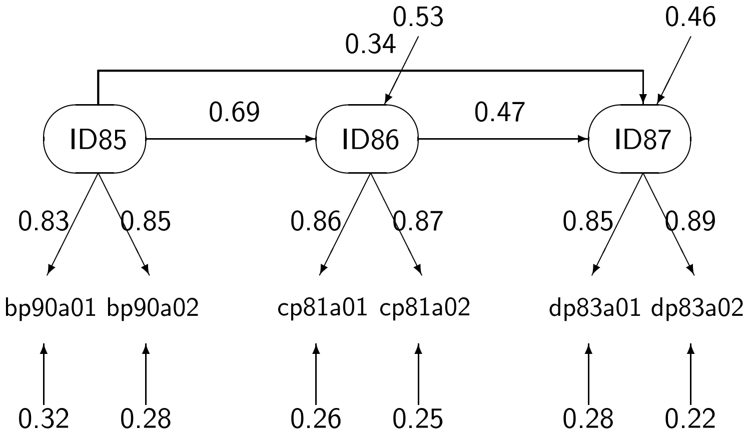
Varianzen, Kovarianzen und Korrelationen der gemessenen Variablen aus dem SOEP für die Jahre 1985, 1986, 1987

	bp90a01	bp90a02	cp81a01	cp81a02	dp83a01	dp83a02
bp90a01 (y_{11})	1.248	0.702	0.559	0.488	0.517	0.484
bp90a02 (y_{21})	0.881	1.068	0.509	0.544	0.467	0.501
cp81a01 (y_{12})	0.686	0.577	1.205	0.745	0.568	0.532
cp81a02 (y_{22})	0.563	0.580	0.844	1.064	0.514	0.529
dp83a01 (y_{13})	0.629	0.525	0.679	0.578	1.186	0.753
dp83a02 (y_{23})	0.537	0.514	0.580	0.542	0.815	0.986

Modellvergleiche durch den χ^2-Differenzentest für das Markov-Modell

Modell	χ^2	df	χ^2_{Diff}	df_{Diff}	RMSEA
Basismodell λs fixiert ϵs fixiert	146.25	12	$--$	$--$	0.081
Variante I λs fixiert ϵs fixiert Autokov.	37.64	8	108.61	4	0.045
Variante 2 λs fixiert Autokov.	2.37	4	35.27	4	0.000
Variante 3 Autokov.	1.92	2	0.45	2	0.000

Ergebnis des Markov-Modells zweiter Ordnung (Variante 2)



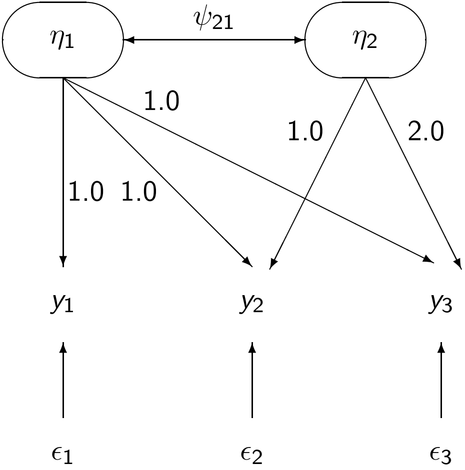
Autokorrelationen der Meßfehler der gemessenen Variablen (Variante 2)

	bp90a01	bp90a02	cp81a01
cp81a01	0.07		
cp81a02		0.04	
dp83a01	0.06		0.06

Strukturgleichungsmodell als Wachstumsmodell

- In einem autoregressiven Panelmodell (formuliert als Kovarianzstrukturgleichungsmodell) werden die Mittelwerte in der Regel auf Null fixiert und nicht weiter berücksichtigt.
- Ohne die Berücksichtigung der Mittelwertsinformationen über die Zeit ist eine Differenzierung nach individuellen und zeitbezogenen Entwicklungstendenzen in Panelmodellen nicht möglich.
- Für die Analyse von Entwicklungsprozessen über die Zeit sind Modellierungen notwendig, die nicht nur individuelle Entwicklungsparameter berücksichtigen, sondern auch die Variation dieser Entwicklungen in der Untersuchungspopulation.
- Die statistische Modellierung von Entwicklungsprozessen hat mit den sogenannten Wachstumsmodellen (*growth curve models*) in verschiedenen Forschungsbereichen eine starke Verbreitung gefunden.

Zweifaktorielles Wachstumsmodell für drei Meßzeitpunkte



Drei Meßzeitpunkte sind notwendig, um ein überidentifiziertes Wachstumsmodell zu spezifizieren.

Drei Meßgleichungen sind zu formulieren:

$$\begin{array}{rcccccc} y_1 & = & \lambda_{11}\eta_1 & + & \lambda_{12}\eta_2 & + & \epsilon_{y_1} \\ y_2 & = & \lambda_{21}\eta_1 & + & \lambda_{22}\eta_2 & + & \epsilon_{y_2} \\ y_3 & = & \lambda_{31}\eta_1 & + & \lambda_{32}\eta_2 & + & \epsilon_{y_3} \end{array} \tag{46}$$

Zwei Strukturgleichungen sind zu formulieren:

$$\begin{array}{rcccc} \eta_1 & = & \alpha_1 & + & \zeta_1 \\ \eta_2 & = & \alpha_2 & + & \zeta_2 \end{array} \tag{47}$$

Man sollte sich genau überlegen, welche Entwicklungsform den Daten zugrundeliegen könnte und eine möglichst sparsame Modellierung vornehmen.

Vier verschiedene Stabilitätsformen können unterschieden werden:

1. *Parallele Stabilität*: Individuelle Entwicklung liegt vor, ohne daß die individuellen Differenzen über die Zeit variieren.

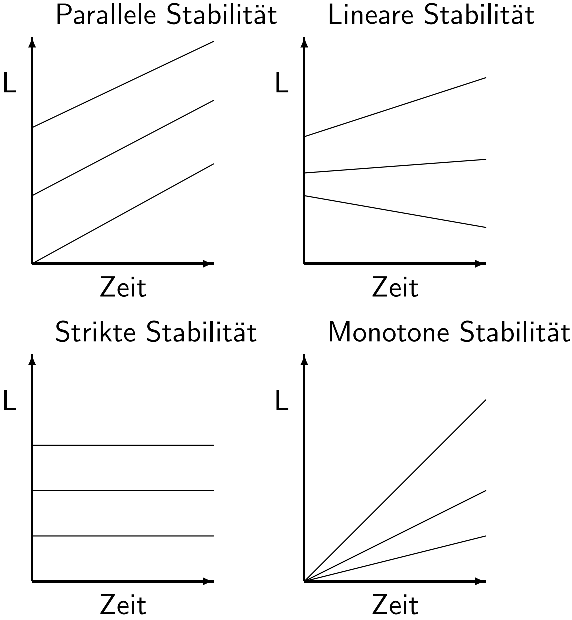
2. *Lineare Stabilität*: Individuelle Entwicklung liegt vor und gleichzeitige Variation der individuellen Differenzen über die Zeit.

3. *Strikte Stabilität*: Es findet keine Entwicklung über die Zeit statt, d. h. $\sigma_{\eta_2} = 0$.

4. *Monotone Stabilität*: Individuellen Differenzen im Wachstum, ohne daß der Anfangsstatus variiert, d. h. $\sigma_{\eta_1} = 0$.

Bei paralleler und linearer Stabilität werden zwei latente Variablen im Wachstumsmodells benötigt.
Bei strikter und monotoner Stabilität wird eine latente Variable im Wachstumsmodell benötigt.

Unterschiedliche Stabilitätsformen in Wachstumsmodellen



L = Level

Beispiel

- Variablen und Daten für ein zweifaktorielles Wachstumsmodell sind der Ausländerstichprobe des SOEP entnommen. Die Stichprobe beinhaltet Ausländer, die der Gruppe der sogenannten Gastarbeiter zugeordnet werden können (Türken, Ex-Jugoslawen, Griechen, Italiener und Spanier).
- Die gemessene Variable y erfaßt das *Zugehörigkeitsgefühl zu Deutschland* (Variablennamen: bp90a01/cp81a01/dp83a01), jeweils abgestuft in fünf Kategorien.
- Die Faktorenladungen sind vollständig restringiert und geben lineares Wachstum wieder:

$$\begin{array}{ccccccc} y_1 & = & 1.0\eta_1 & + & 0.0\eta_2 & + & \epsilon_1 \\ y_2 & = & 1.0\eta_1 & + & 1.0\eta_2 & + & \epsilon_2 \\ y_3 & = & 1.0\eta_1 & + & 2.0\eta_2 & + & \epsilon_3 \end{array} \tag{48}$$

- Die Meßfehler sind jeweils über die Zeit gleichgesetzt ($\epsilon_1 = \epsilon_2 = \epsilon_3$).

Mittelwerte, Varianzen und Kovarianzen der gemessenen Variablen

	$\bar{y}$	bp90a01	cp81a01	dp83a01
bp90a01 (y_1)	3.963	1.248		
cp81a01 (y_2)	3.964	0.686	1.205	
dp83a01 (y_3)	3.953	0.629	0.679	1.168

1. *Basismodell*: Restriktives Meßmodell gemäß Gleichung 48 und gleichen Meßfehlern.
2. *Variante 1*: Ungleiche Meßfehler werden über die Zeit spezifiziert.
3. *Variante 2*: Freisetzung der Faktorenladung für den dritten Meßzeitpunkt (λ_{32}).

Modellvergleiche durch den χ^2-Differenzentest für das zweifaktorielle Wachstumsmodell

Modell	χ^2	df	χ^2_{Diff}	df_{Diff}	RMSEA
Basismodell $\epsilon_1 = \epsilon_2 = \epsilon_3$	3.27	3	$--$	$--$	0.001
Variante I $\epsilon_1 \neq \epsilon_2 \neq \epsilon_3$	0.09	1	3.2	2	0.000
Variante 2 λ_{32} frei	1.56	2	1.7	1	0.000

$$\begin{aligned} \eta_1 &= 3.965 + \zeta_1 \\ \eta_2 &= -.005 + \zeta_2 \end{aligned} \tag{49}$$

Der Mittelwert des Anfangsstatus (η_1) ist von Null verschieden ist ($z = 157.78$). Das Zugehörigkeitsgefühl zu Deutschland (η_2) verändert sich sich aber nicht signifikant über die drei Meßzeitpunkte ($z = -0.40$).

$$\Psi = \begin{pmatrix} 0.743 & \\ & \\ -0.051 & 0.037 \end{pmatrix} \tag{50}$$

Die Varianzen von η_1 ($z = 18.21$) und η_2 ($z = 2.90$) sind von Null verschieden: $\longrightarrow$ keine strikte oder monotone Stabilität.
Die Variation der Wachstumsvariablen ist aber nicht besonders groß.

Zweifaktorielles Wachstumsmodell für die Entwicklung des *Zugehörigkeitsgefühls zu Deutschland*