LISREL-EINFÜHRUNG

Skript zur Lehrveranstaltung

Dr. Lars Satow
FU BERLIN

mit Unterstützung von Seminarteilnehmern

Das Skript ist auch im WWW erhältlich: www.drsatow.de

Inhalt

EINLEITUNG 2

LISREL 2

ANWENDUNGSVORAUSSETZUNGEN 6

FITFUNKTION 7

IDENTIFIKATION 8

DATENAUFBEREITUNG 8

KONFIRMATORISCHE FAKTORENANALYSE 9

MODELLMODIFIKATION 14

EIN EINFACHES STRUKTURGLEICHUNGSMODELL 17

MEHRGRUPPENVERGLEICHE 20

PRELIS 22

FLUSSDIAGAMM 25

LITERATUR 26

Bitte wie folgt zitieren

Satow, L. (2012). LISREL-Einführung: Skript zur Lehrveranstaltung. Freie Universität Berlin. URL: http://www.drsatow.de/lisrel/

EINLEITUNG

Das Skript ist begleitend zum Seminar *Strukturgleichungsmodelle in der Pädagogischen Psychologie* gedacht. Es setzt Grundkenntnisse in der Berechnung von Korrelationen und Kovarianzen voraus. Zudem sollten die Leser über ein hinlängliches Verständnis der linearen Regressionsanalyse und der Faktorenanalyse verfügen. Hier bietet sich die Lektüre der Statistikbücher *Statistik für Sozialwissenschaftler* (Bortz) und *Forschungsmethoden und Evaluation für Sozialwissenschaftler* (Bortz & Döring) an. Hilfreich sind gute EDV-Kenntnisse, insbesondere SPSS-Kenntnisse und der Umgang mit Texteditoren.

Strukturgleichungsmodelle bieten als multivariate Forschungsmethode die Möglichkeit komplexer Datenanalysen. Sie verbinden konfirmatorische Faktorenanalysen mit der linearen Regression und erlauben so die Analyse latenter Strukturen. Hypothetische Konstrukte werden in diesem Ansatz als latente Variablen aufgefaßt. Sie werden mittels mehrerer Indikatoren operationalisiert. So ist es möglich, die Meßfehler der einzelnen Indikatoren zu bestimmen und die „fehlerfreien" regressiven Beziehungen zwischen den hypothetischen Konstrukten zu analysieren, was irreführend auch als Kausalanalyse bezeichnet wird. Strukturgleichungsmodelle werden in lineare Gleichungssysteme umgesetzt. Die unbekannten Parameter können unter bestimmten Voraussetzungen aus den beobachteten Daten geschätzt werden. Eine besondere Rolle spielt dabei die Maximum-Likelihood-Methode, die gleichzeitig die globale Überprüfung eines Strukturgleichungsmodells anhand eines χ^2-Wertes ermöglicht, der sich aus dem Minimum der Fitfunktion ergibt (für eine schöne Darstellung der ML-Methode siehe Revenstorf, 1980).

LISREL

LISREL (Jöreskog & Sörbom, 1989, 1993) ist das wohl bekannteste und verbreitetste Programm, um die Parameter eines Strukturgleichungsmodells zu schätzen. Andere Programme mit ähnlichen Optionen und Möglichkeiten sind EQS, AMOS oder Mx. Wie die Variablen gemessen werden (konfirmatorische Faktorenanalyse) und in welcher Beziehung sie zueinander stehen (Regression), wird in der LISREL-Syntax festgelegt. Es können Meßfehler berücksichtigt und so die Beziehungen zwischen den messfehlerfreien Variablen untersucht werden. LISREL bietet weiterhin die Möglichkeit, ein Modell an verschiedenen Untersuchungsgruppen zu evaluieren (Gruppenvergleich). Durch einen Gruppenvergleich kann z.B. der Einfluß von Moderatorvariablen wie Geschlecht oder Alter untersucht werden.

LISREL verwendet bestimmte Regeln zur Kennzeichnung der Variablen und Matrizen. Griechische Buchstaben kennzeichnen *wahre* Populationsparameter, latente Variablen oder Zufallsfehler. Ein ^ bezeichnet einen Paramterschätzer. Matrizen werden durch Großbuchstaben, Vektoren durch Kleinbuchstaben dargestellt.

Variablen

x, y	Indikatoren, manifeste Variablen (beobachtet)
ξ (Ksi), η (Eta)	Latente Variablen (operationalisiert über Indikatoren)
ζ (zeta), δ (delta), ε (epsilon)	Fehler- bzw. unaufgeklärte Anteile

Die beobachtbaren Variablen werden als Indikatoren oder als *manifeste Variablen* bezeichnet. Im LISREL-Ansatz unterscheidet man zwischen X- und Y- Indikatoren. Erstere stellen Operationalisierungen *exogener* latenter Variablen dar und letztere Operationalisierungen *endogener* latenter Variablen. Exogene Variablen werden im LISREL-Ansatz als KSI-Variablen bezeichnet. Sie stellen unabhängige Variablen dar. Endogene Variablen (ETA) können sowohl die Rolle abhängiger als auch unabhängiger Variablen einnehmen.

Zur Beschreibung der Beziehung zwischen Indikatoren und latenten Variablen werden Matrizen herangezogen. Ein Strukturgleichungsmodell wird durch diese Matrizen abgebildet. Die Matrizen sind wie folgt benannt:

Matrizen als griechische Symbole

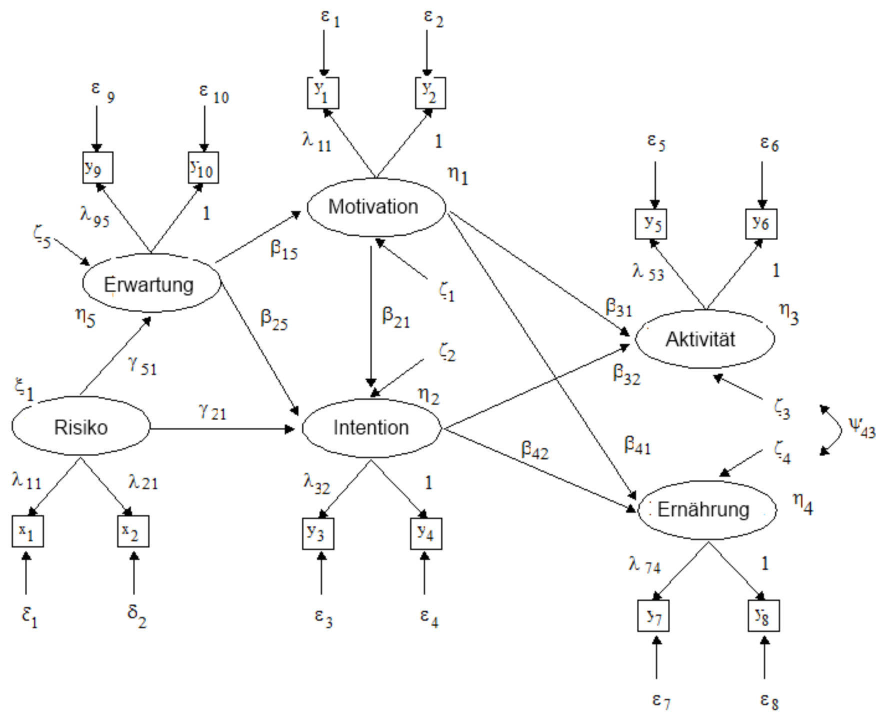
Λx / Λy (gr. lambda)	Faktorladungen
Β (gr. beta), Γ (gr. gamma)	Strukturparameter
Φ (gr. phi), ψ (gr. psi)	Kovarianzmatrizen
Θ (theta)	Fehler-Kovarianz-Matrizen

Matrix	Lisrel-Name	Ordnung
Lambda-Y	LY	NY * NE
Lambda-X	LX	NX * NK
Beta	BE	NE * NE
Gamma	GA	NE * NK
Phi	PH	NK * NK
Psi	PS	NE * NE
Theta-Epsilon	TE	NY * NY
Theta-Delta	TD	NX * NX

(Quelle: Jöreskog & Sörbom, 1989)

Die Spalte „Ordnung“ beschreibt die Spalten und Zeilen einer Matrix. NY steht für *Number of Y-Indicators* etc. So finden sich in den Zeilen der Lambda-Y-Matrix die *Y-Indikatoren*, in den Spalten die *Eta-Variablen*.

Abbildung 1: Strukturgleichungsmodell des Ernährungsverhaltens (aus Satow & Schwarzer, 1997)



Der oben stehenden Abbildung 1 können die Bezeichnungen für die einzelnen Parameter entnommen werden. Die Abbildung erscheint auf den ersten Blick kompliziert. Hilfreich ist daher die Unterscheidung zwischen Indikatoren, die als Kästchen dargestellt werden, und latenten Variablen in Elipsenform. Pfeile zwischen latenten Variablen (β und γ) repräsentieren die Annahme, daß eine latente Variable durch eine andere vorhergesagt wird. Doppelpfeile stehen für Korrelationen bzw. Kovarianzen. Pfeile, die von latenten Variablen auf Indikatoren weisen, veranschaulichen die Meßmodelle. Die Ladungen werden mit λ bezeichnet. Pfeile, die aus dem Nichts auf eine latente Variable weisen, stellen die nicht erklärten Varianzanteile eines Indikators oder einer latenten Variable dar. Diese Varianzanteile werden bei den X-Indikatoren mit δ (delta), bei den Y-Indikatoren mit ε (epsilon) und bei den ETA-Variablen mit ζ (zeta) bezeichnet.

Das Modell weist eine exogene KSI-Variable auf (Risiko) und fünf endogene ETA-Variablen, von denen *Erwartung*, *Motivation* und *Intention* (Swe.) Mediator-Variablen darstellen. Die ETA-Variablen weisen alle eine Restvarianz (ZETA) auf. Die Restvarianzen der beiden abhängigen Variablen sind korreliert (PSI_{43}). Jede latente Variable wird über zwei Indikatoren (immer kleine Kästchen) operationalisiert. Die Pfeile von latenten Variablen auf Indikatoren bringen die Annahme zum Ausdruck, daß eine latente Variable (als hypothetisches Konstrukt) die Ausprägung eines Indikators bestimmt.

Ein solches Strukturgleichungsmodell wird von LISREL als lineares Gleichungssystem aufgefaßt. Die Unbekannten sind die Pfadkoeffizienten, Ladungen, Meßfehlervarianzen etc.

Die Hauptfunktionalität eines Programms zur Analyse von Strukturgleichungsmodellen besteht nun darin, diese Unbekannten zu bestimmen.

Das Gleichungsystem der latenten ETA-Variablen:

$\eta = B\eta + \Gamma\xi + \zeta$,

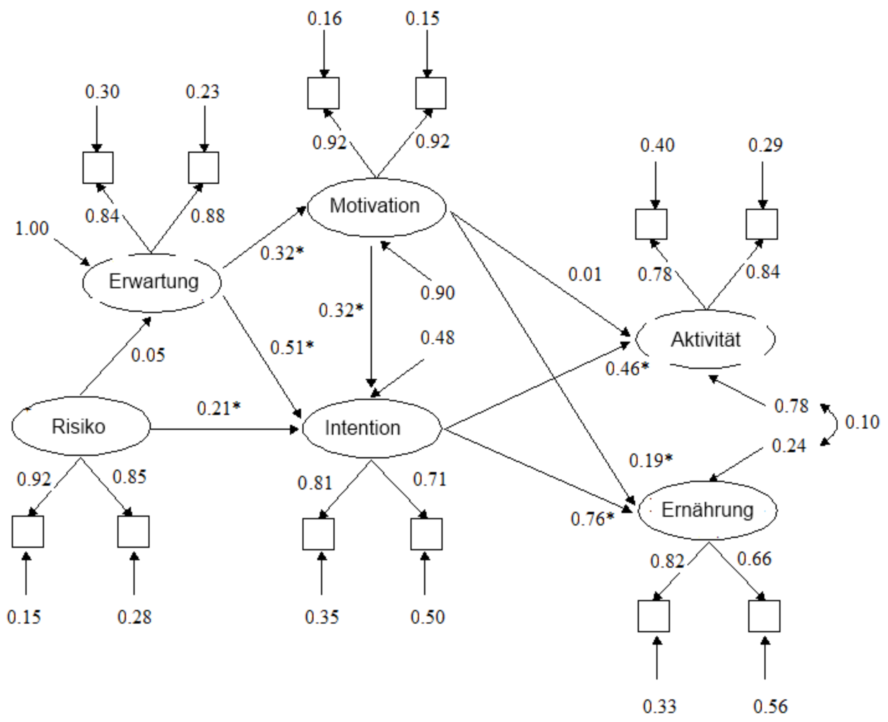
wobei B und Γ Koeffizientenmatrizen sind, deren Elemente die direkten Effekt der ETA-Variablen auf andere ETA-Variablen (B) und der KSI-Variablen auf ETA-Variablen (Γ) repräsentieren.

LISREL kann über die *übliche Matrix-Syntax* oder über eine vereinfachte Syntax (SIMPLIS) programmiert werden. Neuere Versionen erlauben eine Programmierung über ein graphisches Interface: Modelle werden wie in einem Grafik-Programm gezeichnet. Im folgenden wird die zwar komplexe doch sehr flexible klassische Syntax vorgestellt. Sie basiert im wesentlichen darauf, daß ein Strukturgleichungsmodell durch Freisetzen von Zellen bestimmter Matrizen abgebildet wird. Das Freisetzen eines Parameters veranlaßt, daß LISREL den Wert dieses Parameters aus den beobachteten Daten schätzt. Die Indizierung der Zellen und Parameter erfolgt nach der Regel: **Erst Ziel, dann Quelle**. Beispiel: Gegeben sind drei endogene ETA-Variablen, deren Beziehung in der BE(TA)-Matrix abgebildet wird. Soll die erste ETA-Variable durch die dritte ETA-Variable erklärt bzw. vorhergesagt werden (ein Pfeil weist von ETA3 auf ETA1), so wird die Zelle 1/3 der β-Matrix freigesetzt. Dies geschieht z.B. durch den Befehl FREE BE 1 3. In der LISREL-Syntax kann dieser Befehl zu FR BE 1 3 abgekürzt werden.

Nachdem das Modell über die Syntax oder die graphische Schnittstelle spezifiziert wurde, schätzt LISREL die Modellparameter. Dazu verwendet LISREL ein iteratives Verfahren. Zunächst werden Startwerte geschätzt. Ausgehend von diesen Startwerten nähert sich LISREL einer Lösung an. Dieser Prozeß kommt nicht in jedem Fall zu einer Lösung. Spezifikationsfehler, Fehler bei der Dateneingabe, Syntaxfehler u.v.m. können dazu führen, daß LISREL nicht in der Lage ist, die Modellparameter zu bestimmen.

Die Abbildung 2 gibt die Lösung für das in Abbildung 1 dargestellte Strukturgleichungsmodell wieder. Die Pfadkoeffizienten wurden mit LISREL unter Verwendung der Maximum-Likelihood-Funktion geschätzt. Wenn alle Voraussetzungen erfüllt sind, stellen die Parameterschätzungen effiziente Schätzungen der Populationsparameter dar, die durch einen Signifikanztest abgesichert werden können: LISREL berechnet für jeden Parameter einen *t*-Wert der zur Überprüfung der Null-Hypothese (H_0: Parameter = 0) herangezogen werden kann. Der Abbildung 2 ist zu entnehmen, daß die latente Variable *Risiko* signifikant die *Intention* vorhersagt. Diese wird zudem signifikant von den Variablen *Erwartung* und *Motivation* vorhergesagt. *Intention* und *Motivation* erklären *Ernährung* und *Aktivität*.

Abbildung 2: Maximum-Likelihhod-Estimates. Vollständig standardisierte Lösung (aus Satow & Schwarzer, 1997).



ANWENDUNGSVORAUSSETZUNGEN

Die Analysen von Strukturgleichungsmodellen stellt hohe Anforderung an die Daten. Die Anforderungen sind zudem abhängig von der gewählten Methode zur Bestimmung der unbekannten Parameter. Die Annahmen lassen sich in generelle Bedingungen und statistische Bedingungen unterscheiden. Zu den *generellen Bedingungen* zählen die Annahmen, daß die Beziehungen zwischen den Variablen linear, die Effekte der erklärenden auf die abhängigen Variablen additiv, die Beziehungen zwischen den Variablen stochastisch, die gemessenen Variablen kontinuierlich und intervallskaliert sind sowie, daß die Daten durch den Mittelwert, die Varianz und die Kovarianz der gemessenen Variablen repräsentiert werden können (multivariate Normalverteilung) (vgl. Ecob & Cuttance, 1987).

Die wichtigsten Voraussetzungen betreffen die multivariate Normalverteilung der Indikatoren. Das Verfahren liefert aber auch bei *moderater* Abweichung robuste Schätzungen der Parameter. Darüber hinaus sollte die Stichprobe hinreichend groß sein. In der Literatur wird eine Stichprobengröße von mind. N = 200 bis N = 400 Personen nahegelegt. Die Indikatoren sollten kontinuierlich bzw. intervallskaliert sein. Ordinale Indikatoren können jedoch mittels bestimmter Korrelationstechniken ebenfalls verwendet werden. Bevor man mit der Datenanalyse beginnt, sollte man zumindest visuell überprüfen, ob die Daten annähernd normalverteilt sind, und überlegen, ob die Annahme eines Intervallskalenniveaus gerechtfertigt erscheint.

FITFUNKTION

Mit Hilfe der Fitfunktion ist LISREL in der Lage, die Lösung eines linearen Strukturgleichungsmodells zu finden und die gesuchten Parameter zu bestimmen. Die „beste" Fitfunktion ist das Maximum-Likelihood-Verfahren. Sie optimiert die Wahrscheinlichkeit für das Auftreten der beobachteten Daten, indem die Modellparameter so bestimmt werden, daß sie mit größter Wahrscheinlichkeit die Stichprobenmatrix verursacht haben.

$$L(\Theta) = \log / \Sigma / + \mathrm{trace}(S\Sigma^{-1}) - \log / S / - p$$

ML-Schätzungen haben die folgenden asymptotischen (large-sample) Eigenschaften:

a) Sie sind konsistent

b) Sie haben eine minimale Cramer-Rao-Stichprobenvarianz

c) Sie sind multivariat normalverteilt

Zudem kann aus dem Wert der ML-Fitfunktion ein χ^2-Wert berechnet werden. Dieser erlaubt die statistische Evaluation eines Modells. Ein signifikanter χ^2-Wert bedeutet, daß das spezifizierte Modell die Struktur auf Populationsebene nicht exakt repräsentiert. Der χ^2-Wert sollte daher möglichst kein ausfallen und nicht signifikant werden. Auf die Probleme der Anwendung des χ^2-Tests wird später detailliert eingegangen.

$$\chi^2 = (N-1)\, ML(\Theta)$$

Die Freiheitsgrade des χ^2-Wertes berechnen sich nach

$$df = p\,(p+1)/2-q,$$

wobei

p: Anzahl der Variablen in der Stichprobenmatrix,

q: Anzahl der gesuchten Parameter,

(Bentler & Bonett, 1980).

Vorteile der ML-Fitfunktion

- Der Wert der ML-Fitfunktion ist χ^2 verteilt (für große Stichproben). Er kann inferenzstatistisch abgesichert werden.

- die ML-Fitfunktion ist sehr effektiv. Sie liefert die präzisesten Parameterschätzungen.

Nachteile

- die ML-Fitfunktion setzt eine multivariate Normalverteilung der Variablen voraus.
- sie ist nur für große Stichproben sinnvoll (N > 200)

IDENTIFIKATION

Die Frage, ob eine Lösung mathematisch eindeutig ist, wird als *Identifikationsproblem* bezeichnet. Eine *notwendige* Bedingung für die Identifikation der Lösung ist eine positive Anzahl von Freiheitsgraden. Diese errechnen sich aus den vorgegebenen Kovarianzen und der Anzahl gesuchter Parameter (s.o.). Aber auch bei einer positiven Anzahl von Freiheitsgraden ist es nicht immer möglich zu entscheiden, ob die gefundene Lösung mathematisch eindeutig ist.

LISREL verwendet eine Heuristik, um Identifikationsprobleme aufzuspüren. Auf ein Identifikationsproblem deutet die LISREL-Warnung „matrix not positive definite“ hin. Weitere Hinweise auf ein Identifikationsproblem sind sehr auffällige, unbegründbare Schätzungen und mathematisch nicht mögliche Werte (negative Varianzen).

Es gibt keine standardisierte Methode die Identifizierung zu prüfen. Eine Möglichkeit besteht darin, unterschiedliche *Starting values* zu definieren. Von diesen ausgehend schätzt LISREL in einem iterativen Prozeß die endgültigen Koeffizienten. Ein Identifikationsproblem ist wahrscheinlich, wenn LISREL bei unterschiedlichen Starting values zu unterschiedlichen Koeffizienten gelangt. Ein anderer Test besteht darin, den Parametern eines Modells Werte zuzuweisen, die leicht von den LISREL-Schätzungen abweichen. Ein Identifikationsproblem ist dann wahrscheinlich, wenn sich der Modellfit dadurch nicht deutlich verschlechtert (Hayduk, 1987).

DATENAUFBEREITUNG

Den Ausgangspunkt der LISREL-Analysen bildet eine Stichprobenmatrix. Diese kann (a) entweder aus Rohwerten direkt von LISREL berechnet oder (b) vor den LISREL-Analysen mit PRELIS erzeugt werden. Als dritte Möglichkeit kann die Matrix auch direkt in die LISREL-Syntax integriert werden. In der Regel sollte eine *Kovarianzmatrix* verwendet werden, da der Ansatz für die Analyse von Kovarianzstrukturen entwickelt wurde (Jöreskog & Sörbom, 1989; Hair, 1995). Dies ist vor allem dann sinnvoll, wenn die Einheiten der Messung bedeutungsvoll sind. Die Verwendung von *Korrelationsmatrizen* kann zu verzerrten Fitstatistiken und Standardfehlern führen.

Im Folgenden wird an einem Beispiel erläutert, wie SPSS-Daten für LISREL aufbereitet werden. Zunächst werden fehlende Daten durch einen Platzhalter (z.B. 99) gekennzeichnet oder Fälle mit fehlenden Daten ausgeschlossen. **In keinem Fall darf das Datenfile Leerzeichen zur Kennzeichnung fehlender Daten enthalten**. Im zweiten Schritt werden die Variablen so umformatiert, daß ein *Punkt* als Dezimalzeichen verwendet wird. Schließlich werden die Rohdaten in ein ASCII-File geschrieben. Ein ASCII-File ist ein einfaches Text-File, das mit jedem Textverarbeitungsprogramm bearbeitet werden kann.

In diesem Beispiel werden die Daten für zwei Indikatoren aufbereitet. Bei den Indikatoren handelt es sich um zwei Tests (WOR & EMO), die Ängstlichkeit von Schülern erfassen. Die folgenden SPSS-Befehle sind notwendig, um aus einem SPSS-Datenfile ein ASCII-File für LISREL herzustellen.

Der erste SPSS-Befehl löscht Fälle mit fehlenden Daten:

```
FILTER OFF.
USE ALL.
SELECT IF( ~missing(wor) and  ~missing(emo)).
EXECUTE .
```

Möchte man Fälle mit fehlenden Daten im Datensatz behalten, ist es **unbedingt** erforderlich, Platzhalter (hier 99) für fehlende Daten einzuführen:

```
RECODE
  emo wor  (MISSING=99)  .
EXECUTE .
```

Der nächste SPSS-Befehl ersetzt das Komma durch einen Punkt als Dezimalzeichen:

```
formats
wor emo(COMMA6.2) .
exec.
```

Dieser Befehl schreibt die Daten in ein ASCII-Text-File namens TEST.DAT. Die Reihenfolge der Variablen in der Text-Datei entspricht der Reihenfolge im SPSS-Datenfile!

```
WRITE OUTFILE='C:\Eigene Dateien\test.dat'
  TABLE /ALL.
EXECUTE.
```

Eine weitere Aufbereitung der Daten kann mit PRELIS geschehen. So bietet PRELIS verschiedene Möglichkeiten zum Umgang mit fehlenden Daten und erlaubt darüber hinaus die Berechnung von Korrelationen für ordinale Variablen.

KONFIRMATORISCHE FAKTORENANALYSE

Mittels einer konfirmatorischen Faktorenanalyse kann überprüft werden, ob Variablenausprägungen (z.B. Rechenaufgaben) bestimmte Faktoren (z.B. Intelligenz) zugrunde liegen. Die Anzahl der Faktoren wird dabei im Gegensatz zum Vorgehen bei einer explorativen Faktorenanalyse vorab festgelegt. Ebenso wird vorab bestimmt, welche Indikatoren mit welchen Faktoren in Zusammenhang stehen. Ob die getroffenen Annahmen geeignet sind, die Datenlage zu beschreiben, kann dann anhand der Fitindizes beurteilt werden. Konfirmatorische Faktorenanalysen eignen sich damit besonders zur Überprüfung von fundierten theoretischen Überlegungen.

Im folgenden wird mit LISREL überprüft, ob den acht Items eines Tests für Schüler, der Selbstwirksamkeitserwartungen messen soll, *ein* Faktor zugrunde liegt. Inhaltlich wird damit überprüft, ob das hypothetische Konstrukt *Selbstwirksamkeitserwartungen*, das der Test erfassen soll, die Antworten auf die Items determiniert. Als Methode zur Ermittlung der

Ladungen wird das Maximum-Likelihood-Verfahren gewählt. Hier werden die Populationsparameter so bestimmt, daß die Wahrscheinlichkeit für das Auftreten der beobachteten Stichprobendaten maximiert wird. Das ML-Verfahren trifft jedoch die erwähnten Annahmen, darunter die einer hinreichend großen Stichprobe und einer multivariaten Normalverteilung der Indikatoren.

Die LISREL-SYNTAX lautet:

```
DA NI=8 NO=2691 MI=99 MA=CM
LA
item1 item2 item3 item4 item5 item6 item7 item8/
RA FI=daten.dat
MO NX=8 NK=1 LX=FU,FI PH=ST TD=DI,FR
LK
Faktor1/
FR LX 1 1 LX 2 1 LX 3 1 LX 4 1 LX 5 1 LX 6 1 LX 7 1 LX 8 1
OU ML SC
```

Die Spezifikation beginnt mit dem Kennwort DATA, welches zu DA abgekürzt werden kann. Generell können alle Syntaxbefehle auf die *ersten beiden Buchstaben* gekürzt werden. In der Data-Zeile wird die Anzahl der Indikatoren (NI=8; NI steht für number of indicators) sowie die Anzahl der Versuchspersonen (NO=2691, NO steht für number of oberservations) angegeben. Mit MI(SSING)=99 wird der Platzhalter für fehlende Daten festgelegt. Zum Schluß wird mit MA=CM bestimmt, daß aus den Rohdaten, die in der Zeile „RA FI = daten.dat" eingelesen werden, eine Kovarianzmatrix berechnet wird (CM). Alternativ könnte hier z.B. auch eine Korrelationsmatrix (KM) angefordert werden. Diese Matrix bildet dann den Ausgangspunkt für die Schätzung der Parameter. Das Kennwort LA für LABEL leitet die Bezeichnung der Indikatoren ein. Sie ist willkürlich. Die Bezeichnung der Indikatoren schließt mit einem *Slash* ab.

Das Kernstück bildet die MO-Zeile. Hier wird festgelegt, welche Matrizen verwendet werden und welche Form sie aufweisen. Es sind nicht immer alle Zellen einer Matrix notwendig, um ein Modell abzubilden. So interessiert i.R. nur die Hauptdiagonale der Theta-Epsilon-Matix und der Theta-Delta-Matrix und seltener die übrigen Zellen dieser Matrizen. Mit Hilfe dieser Hauptdiagonalen werden die Meßfehlervarianzen abgebildet. Die LISREL-Syntax erlaubt es, Matrizen auf eine bestimmte Form festzulegen. So bedeutet der Befehl TD=DI,FR, daß die Theta-Delta-Matrix auf die Hauptdiagonale (DIAGONAL) reduziert wird. Mit FR(EE) wird diese Hauptdiagonale freigesetzt, was in der LISREL-Syntax bedeutet, daß die freigesetzten Parameter von LISREL aus den beobachteten Daten geschätzt werden sollen. In diesem Fall sollen also die Meßfehler der X-Indikatoren von LISREL geschätzt werden. Zuvor wird jedoch die Anzahl der X-Indikatoren NX=8 und die Anzahl der exogenen Faktoren (NK=1) festgelegt. In der λx-Matrix (LX) spezifiziern wir, welche X-Indikatoren durch welche Faktoren erklärt werden. Die PH-Matrix enthält die Varianzen und Kovarianzen der latenten Faktoren. Durch PH=ST wird sie standardisiert, d.h. die Hauptdiagonale wird auf eins festgelegt. Durch diese Standardisierung wird den latenten KSI-Faktoren eine Varianz von 1 zugewiesen.

Mit LK wird die Bezeichnung der latenten Faktoren eingeleitet. Im nächsten Absatz wird durch das Kennwort FR (FREE) die Freisetzung bestimmter Zellen bestimmter Matrizen vorgenommen. Die Freisetzung einer Zelle bewirkt, daß der entsprechende Parameter von

LISREL geschätzt wird. In diesem Beispiel werden die Zellen der λx-Matrix so freigesetzt, daß die acht X-Indikatoren durch den Faktor erklärt werden. Beispiel: Das zweite Item wird durch den ersten Faktor erklärt. Diese Beziehung wird in der λx-Matrix durch das Freisetzen der Zelle 2 / 1 (Ziel / Quelle) abgebildet.

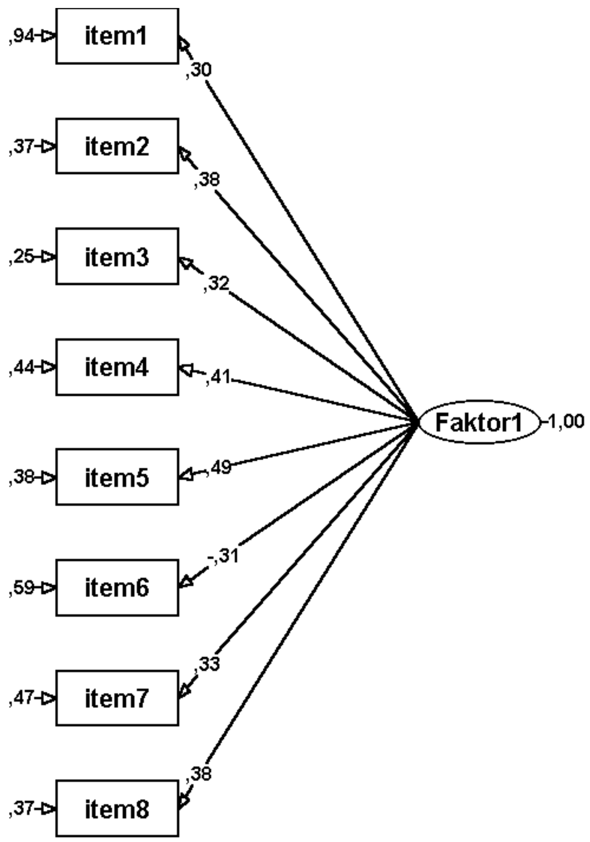
Die OU-Zeile schließt die Syntax ab. In ihr wird zum einen die Methode zur Schätzung der Modellparameter festgelegt (hier ML). Zum anderen kann der LISREL-Output näher spezifiziert werden (hier wird mit SC eine vollständig standardisierte Lösung angefordert).

In der vereinfachten Syntax SIMPLIS schreiben wir dafür folgenden LISREL-Job:

```
Konfirmatorische Faktorenanalyse mit 8 Items
Observed Variables
  item1 item2 item3 item4 item5 item6 item7 item8
Covariance From File MATRIX.CM
Sample Size 2691
Latent Variables: FAKTOR1
Relationships:
   item1 - item8 = FAKTOR1
End of Problem
```

Die Abbildung 3 zeigt die Maximum-Likelihood-Lösung. Da die Varianz der Indikatoren bzw. Items nicht standardisiert ist, ergibt sich die Varianzaufklärung für die einzelnen Indikatoren nicht unmittelbar aus dem Quadrat der Ladung: Die Größe der Ladungen muß in Relation zur Varianz der Items interpretiert werden. Einfacher ist dies in der *vollständig standardisierten* Lösung. Hier ist die Varianz der Items auf 1 umgerechnet.

Abbildung 3: Maximum-Likelihood-Estimates



Der Fit eines Modells läßt sich anhand globaler Fitindizes beurteilen. Häufig verwendet werden dazu der χ^2-Wert, das Verhältnis von χ^2-Wert zu der Anzahl der Freiheitsgrade, der GFI sowie der RMR (Tabelle 1).

Tabelle 1: Globale Fitindizes

χ^2	df	p	GFI	NFI	SRMR
181	20	.00	.98	.93	.034

Für die meisten Fitindizes bestehen *keine* konkreten Grenzwerte. Als Daumenregel sollte der χ^2-Wert nicht fünfmal größer als die Freiheitsgrade ausfallen, der GFI sollte größer als .90 und der SRMR kleiner als .050 werden.

Die χ^2 Statistik ermöglicht den statistischen Tests des Modells. Voraussetzung für die Anwendung des χ^2 - Wertes als Teststatistik:

- das Modell ist gültig, alle Annahmen werden erfüllt
- die Stichprobe ist ausreichend groß

- multivariate Normalverteilung
- die ML-Fitfunktion wurde verwendet

Der χ^2 - Wert sollte jedoch weniger als Teststatistik denn viel mehr als ein globaler Fitindex interpretiert werden, der die Differenz zwischen der Stichprobenkovarianzmatrix und der aufgrund des Modells reproduzierten Kovarianzmatrix mißt (Jöreskog, 1993). Generell sollte der χ^2-Wert möglichst klein ausfallen. Für den χ^2 - Test (H0: $\chi^2 = 0$) wird davon ausgegangen, daß das Modell exakt die Zusammenhänge auf Populationsebene beschreibt (Jöreskog, 1993, S. 309). Diesem Anspruch werden Modelle jedoch häufig nicht gerecht. Kleine Abweichungen führen daher bei einer großen Stichprobe zu einem signifikanten χ^2 - Wert. Generell gilt: Je größer die Stichprobe, desto größer wird die Wahrscheinlichkeit, daß ein signifikanter Unterschied entdeckt wird. Die Teststärke (statistical power) wächst mit der Stichprobengröße an, so daß das zu prüfende Modell möglicherweise abgelehnt wird, obwohl es nur *geringfügig* vom tatsächlichen Modell abweicht. Dies ist besonders in den Sozialwissenschaften relevant, wo es nur selten Modellvorstellungen gibt, die den Anspruch haben, die Realität *exakt* zu beschreiben (Bentler & Bonett, 1980).

Geschachtelte Modelle, die sich nur in der Anzahl der gesuchten Parameter unterscheiden, können anhand der χ^2-Werte-Differenz verglichen werden. Die Differenz der χ^2-Werte der beiden Modelle ist ebenfalls χ^2-verteilt. Die Anzahl der Freiheitsgrade entspricht der Differenz der Freiheitsgrade der beiden Modelle (Hox, 1995). Je größer die Differenz der χ^2-Werte im Vergleich zu den Freiheitsgraden, desto plausibler ist die Annahme eines spezifischeren Modells (Wald-Test).

Der GFI wurde von Jöreskog & Sörbom eingeführt. Der GFI kann als ein genereller *Determinationskoeffizient* verstanden werden (Tanaka, 1993, S. 19), bevorzugt jedoch komplexe Modelle gegenüber sparsamen Modellen. Der AGFI (adjusted GFI) stellt eine Erweiterung des GFI dar. Der AGFI berücksichtigt die Anzahl der Freiheitsgrade des aktuellen Modells im Vergleich zu der Anzahl der Freiheitsgrade eines Nullmodells (Hair, 1995, S. 686).

Der NFI wurde von Bentler und Bonett eingeführt. Der NFI errechnet sich aus der Differenz zwischen dem χ^2 - Wert des Nullmodells und des aktuellen Modells im Verhältnis zu dem χ^2 - Wert des Nullmodells. Interpretation: Ein Modell befindet sich auf dem Weg vom Nullmodell (NFI = 0) zum perfekten Modell (NFI = 1). Der PNFI (James et al., 1982) berücksichtigt zudem wie der AGFI die Anzahl der Freiheitsgrade des aktuellen Modells gegenüber dem Nullmodell: df(aktuell) / df (null) * NFI.

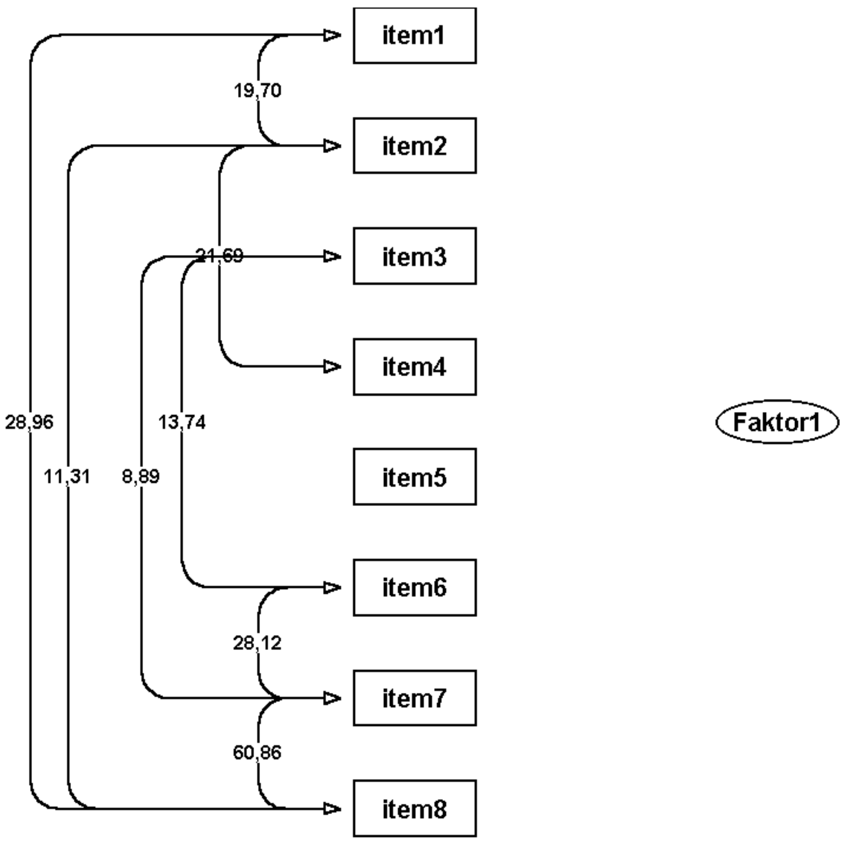
Der RMR ist ein Maß für die Differenz zwischen der Stichprobenmatrix und der aufgrund des Modells reproduzierten Matrix. Der RMR wird als durchschnittliches Residuum nach dem Fitprozeß interpretiert. Der standardisierte SRMR sollte möglichst kleiner als .05 ausfallen.

$$RMR = \sqrt{2\frac{\sum\sum (s_{ij} - \hat{\sigma}_{ij})^2}{(p+q)(p+q+1)}}$$

MODELLMODIFIKATION

Einen Einblick in den detaillierten Modellfit erlauben die *Modification Indizes (MI)* (Abbildung 4). Sie geben an, um welchen Betrag der χ^2-Wert sich verbessert, wenn ein zusätzlicher Parameter geschätzt wird, was der Aufgabe eines Freiheitsgrades gleichkommt. Als Daumenregel weist ein MI > 15 auf eine möglicherweise notwendige Modellmodifikation hin.

Abbildung 4: Modification Indices



Eine besondere Verbesserung des χ^2-Wertes wird von der Freisetzung der Kovarianz der Meßfehler der Items 7 und 8 erwartet. Das Freisetzen dieser Kovarianz würde zu einer Verringerung des χ^2-Wertes von 60.86 Punkten führen. Ist diese Verbesserung signifikant (s. χ^2-Differenztest) kann man in Erwägung ziehen, diesen zusätzlichen Parameter zuzulassen. Die Meßfehlerkovarianz sollte jedoch nur dann freigesetzt werden, wenn dies auch *theoretisch plausibel* erscheint. Standen Item 7 und 8 z.B. am Ende eines langen Tests, dann ist anzunehmen, daß beide durch zunehmende Ermüdung beeinflußt wurden: Ihre Meßfehler könnten aus diesem Grund kovariieren.

Für die vollständig standardisierte Lösung (Tabelle 2) werden alle Varianzen auf 1 umgerechnet, und Pfadkoeffizienten lassen sich im Sinne standardisierter direkter Effekte interpretieren. Durch die Standardisierung, wird deutlich, daß das 1. Items der schwächste Indikator des Faktors ist (Ladung = .29). Indikator und Faktor teilen sich in diesem Fall rund 8% Varianz (.29 * .29 = .084), was einem Meßfehler des Indikators von 92% entspricht. Das 1. Item scheint damit für die Operationalisierung des latenten Konstruktes ungeeignet. Das Vorzeichen für Item Nr. 6 ist negativ, da dieses Item mangelnde Selbstwirksamkeitserwartung mißt.

Tabelle 2: Standardisierte Lösung

```
Konfirmatorische Faktorenanalyse
 COMPLETELY STANDARDIZED SOLUTION

         LAMBDA-X

             Faktor1
            --------
     item1       0.29
     item2       0.53
     item3       0.54
     item4       0.52
     item5       0.62
     item6      -0.38
     item7       0.43
     item8       0.53
```

Aufgrund dieser Analysen wurden zwei Modifikationen vorgenommen. (a) Das Item Nr. 1 wurde eliminiert und (b) die Meßfehlerkovarianz zwischen den Items 7 und 8 wurde zugelassen. Dazu wurde die Syntax leicht modifiziert. Nach SE(LECT) werden die interessierenden Items ausgewählt (ohne Item Nr. 1). Da nun ein Indikator weniger in das Modell eingehen soll, wird in der MO-Zeile die Anzahl der X-Indikatoren mit NX=7 festgelegt. Um die Meßfehlerkovarianzen freizugeben, wird die TD-Matrix zunächst als rechteckig und fixiert spezifiziert (TD=FULL,FIXED). Erst danach werden die Hauptdiagonale und die Zelle 7/6 (=Meßfehlerkovarianz) freigesetzt:

```
Konfirmatorische Faktorenanalyse
DA NG=1 NI=8 NO=2691 MA=CM
LA
item1 item2 item3 item4 item5 item6 item7 item8/
SE
2 3 4 5 6 7 8/
CM FU  FI=matrix.cm
MO NX=7 NK=1 LX=FU,FI PH=ST TD=FU,FI
LK
Faktor1/
FR TD 1 1 TD 2 2 TD 3 3 TD 4 4 TD 5 5 TD 6 6 TD 7 7 TD 7 6
FR LX 1 1 LX 2 1 LX 3 1 LX 4 1 LX 5 1 LX 6 1 LX 7 1
PATH DIAGRAM
OU SC
```

Der Modellfit verbesserte sich gemessen am χ^2-Wert durch diese Modifikationen deutlich. Allerdings weist das modifizierte Modell auch weniger Freiheitsgrade auf. Für beide Modelle wird der χ^2-Wert signifikant. Beide Modelle müßten daher aufgrund des χ^2-Tests verworfen werden. Dieser Test wird aus den oben genannten Gründen hier jedoch nicht angewendet. Der

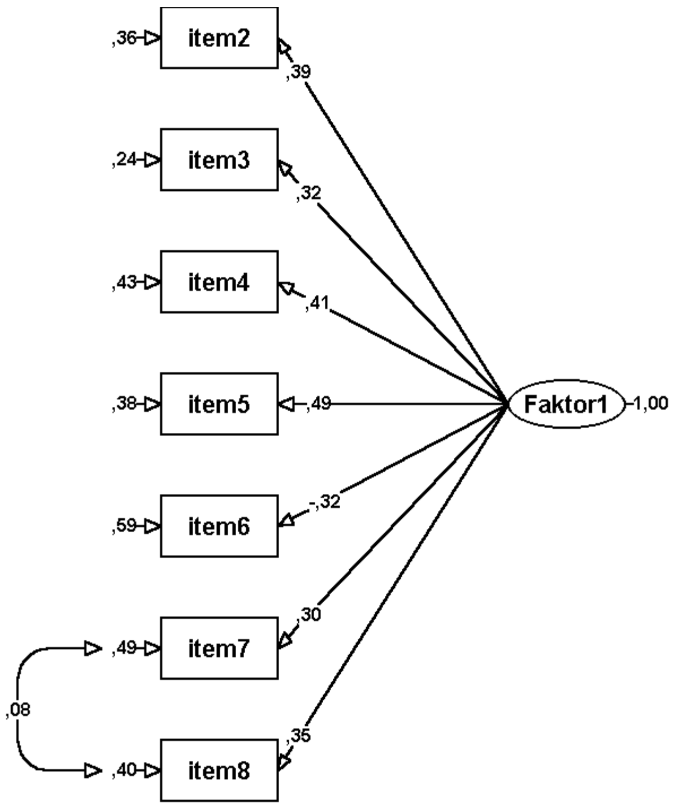
χ^2-Wert wird vielmehr als ein Hinweis auf die Anpassungsgüte interpretiert: Er sollte nicht 5 mal größer als die Anzahl der Freiheitsgrade sein. Die anderen Indizes zeigen einen guten Modellfit an, der sich durch die Modifikation leicht verbesserte (Tabelle 3).

Tabelle 3: Globale Fitindizes

	χ^2	df	p	GFI	NFI	SRMR
Ursprüngliches Modell	181	20	.00	.98	.93	.034
Modifiziertes Modell	66	13	.00	.99	.97	.023

Der Abbildung 5 sind die ML-Schätzungen für das modifizierte Modell zu entnehmen. Dem Modell liegen nunmehr 7 statt 8 Indikatoren zugrunde. Zwischen Item 7 und Item 8 wurde die Meßfehlerkovarianz geschätzt. Aufgrund dieser Modifikation zeigen sich geringe Abweichungen in den Ladungen.

Abbildung 5: Modifiziertes Modell



EIN EINFACHES STRUKTURGLEICHUNGSMODELL

Das Vorgehen zur Analyse eines Strukturgleichungsmodells läßt sich in vier Schritte fassen:

1. Theoretische Überlegungen führen zu einem Strukturgleichungsmodell. Das Modell legt fest, welche latenten Variablen wie operationalisiert werden (Meßmodelle) und in welchem Verhältnis sie zueinander stehen (Strukturmodell).
2. Datenerhebung und -aufbereitung
3. Umsetzung in die LISREL-Syntax. (Die Meßmodelle sollten zunächst getrennt überprüft werden)
4. Schätzung der Modellparameter und des globalen Modellfits.
5. Modifikation des Modells aufgrund theoretischer Überlegungen.

In diesem fiktiven Beispiel geht es um den Zusammenhang zwischen Schulleistung, Prüfungsangst und schulbezogenen Selbstwirksamkeitserwartungen. Alle latenten Variablen werden *in diesem Beispiel* mit zwei Indikatoren operationalisiert: Die *Schulleistung* durch die Zensuren in Deutsch (DNOTE) und Englisch (ENOTE), die *Prüfungsangst* durch Besorgtheit (WORRY) und Aufgeregtheit (EMO), die *Selbstwirksamkeitserwartungen* durch Hilflosigkeit (HILFL) und Selbstwirksamkeitserwartung (SWE). Theoretische Überlegungen führen zu der Annahme, daß die *Schulleistung* (ETA1) durch die *Prüfungsangst* (KSI2) und durch die *Selbstwirksamkeitserwartungen* (KSI1) beeinflußt wird.

Diese Annahmen werden in die LISREL-Syntax umgesetzt. Aus den 7 Indikatoren werden mit SELECT die 6 interessierenden Variablen ausgewählt. Wichtig ist dabei, daß zuerst die Y-Indikatoren (DNOTE & ENOTE) der endogenen Variablen *Leistung* und dann die X-Indikatoren der exogenen Variablen *Selbstwirksamkeitserwartung* und *Prüfungsangst* ausgewählt werden. Entsprechend wird in der MO-Zeile die Anzahl der X-Indikatoren auf NX=4 und die Anzahl der Y-Indikatoren auf NY=2 festgelegt. Die 4 X-Indikatoren sollen als Operationalisierung für 2 KSI-Variablen (NK=2) dienen, die 2 Y-Indikatoren für eine ETA-Variable (NE=1). Die λy-Matrix beschreibt analog zur λx-Matrix, welche Indikatoren mit welchen latenten Variablen in Zusammenhang stehen. Die PSI-Matrix wird durch PS=DI,FR in der Hauptdiagonalen freigesetzt. Dadurch werden die nicht erklärten Varianzanteile der ETA-Variablen geschätzt. Die Theta-Epsilon-Matrix (analog zur Theta-Delta-Matrix) erlaubt die Freigabe der Meßfehlervarianzen und -kovarianzen der Y-Indikatoren. Die γ-Matrix (GA) erlaubt die Spezifikation der Pfade zwischen endogenen und exogenen Variablen.

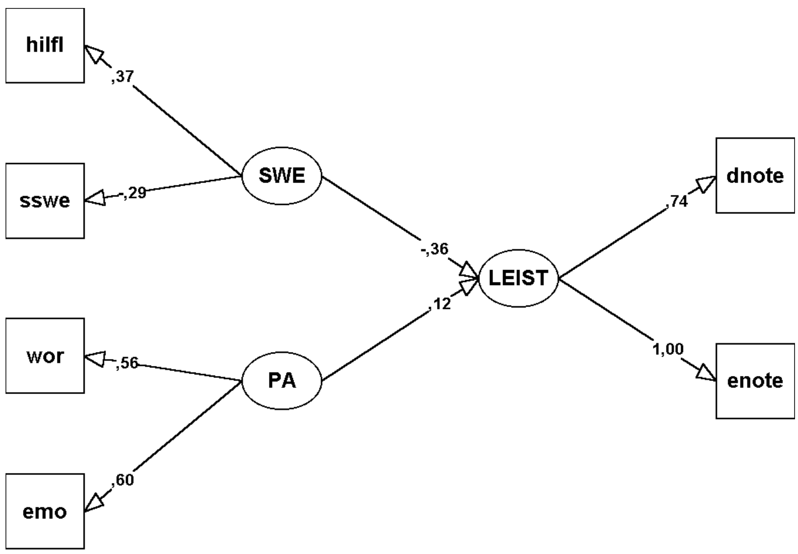
Mit LK und LE werden jeweils die KSI- und die ETA-Variablen benannt. Darauffolgend werden gemäß der Operationalisierung mit FR die einzelnen Parameter der LX- und LY-Matrix zur Schätzung freigegeben. Um die Identifizierbarkeit des Meßmodells für die ETA-Variable zu gewährleisten, wird mit VALUE 1 LY 2 1 eine Ladung auf 1 fixiert. Gleichzeitig wird dadurch die Einheit der ETA-Variable festgelegt. Generell gilt, daß innerhalb eines Meßmodells ein Parameter auf einen bestimmten Wert gesetzt werden muß. Für die KSI-Variablen geschieht das durch PH=ST in der MO-Zeile: Die Varianzen werden auf 1 gesetzt. Schließlich wird die γ-Matrix den theoretischen Annahmen entsprechend freigesetzt.

In der OU-Zeile wird mit SC eine vollständig standardisierte Lösung angefordert. Mit IT=100 wird die Anzahl der Iterationen zur Bestimmung der Lösung auf maximal 100 festgelegt. LISREL führt während des Iterationsprozesses eine Prüfung durch, um festzustellen, ob LISREL sich noch auf dem richtigen Weg befindet. Diese Prüfung führt in einigen Fällen dazu, daß keine Lösung gefunden wird. Sie kann mit AD=OFF abgewählt werden:

```
Strukturgleichungsmodell
DA NI=7 NO=2706 MA=CM
LA
zschul hilfl sswe wor emo dnote enote /
SE
dnote enote hilfl sswe wor emo /
RA FI=data.dat
MO NX=4 NK=2 LX=FU,FI PH=ST TD=DI,FR  c
   NY=2 NE=1 LY=FU,FI PS=DI,FR TE=DI,FR GA=FU,FI
LK
SWE PA/
LE
LEIST/
FR LX 1 1 LX 2 1
FR LX 3 2 LX 4 2
FR LY 1 1
VA 1 LY 2 1
FR GA  1 1 GA 1 2
OU SC IT=100 AD=OF
```

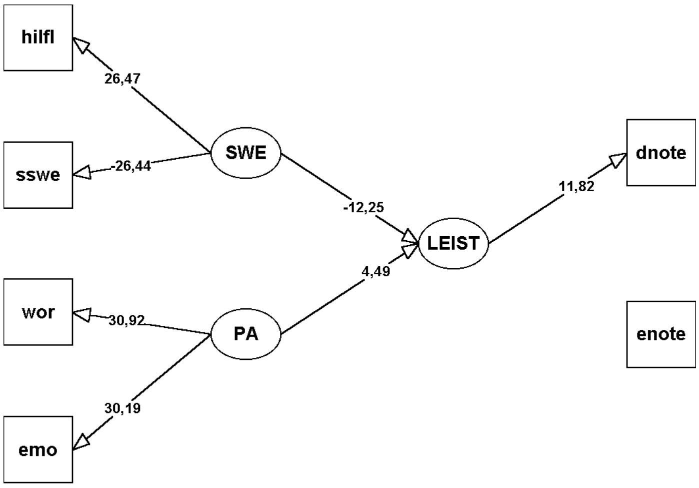
Die ML-Estimates sind in der folgenden Abbildung 6 wiedergegeben. Wie bei der konfirmatorischen Faktorenanalyse lassen sich die Koeffizienten hinsichtlich ihrer Größe nur in Relation zur Varianz der Indikatoren und latenten Variablen richtig interpretieren.

Abbildung 6: Strukturgleichungsmodell



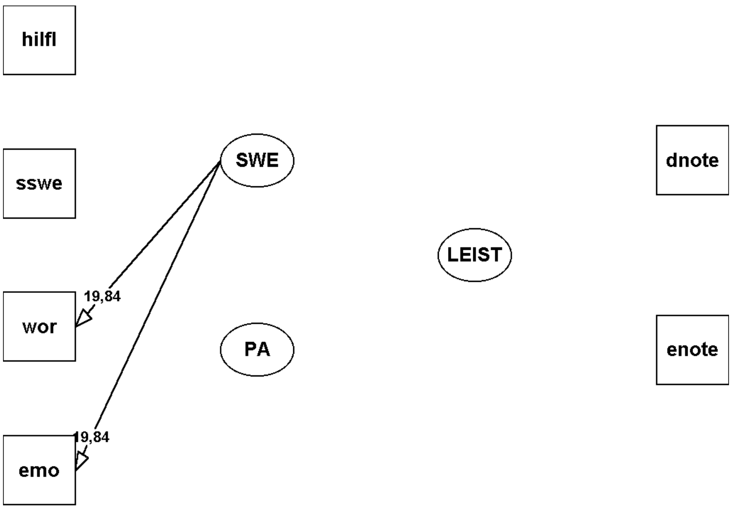
Aufschluß über die Bedeutsamkeit der Koeffizienten geben die *t*-Werte. Im LISREL-Ansatz wird bei t > 2.00 von signifikanten Koeffizienten ausgegangen. Aus der Abbildung 7, in welcher die *t*-Werte der Parameter angezeigt werden, ist ersichtlich, daß alle Koeffizienten hoch signifikant sind (kleinster t-Wert von 4,49 für den Pfad von Prüfungsangst auf Leistung).

Abbildung 7: *t*-Werte



Für den *fixierten* Koeffizienten LY 2 1 wird kein t-Wert angegeben, da er nicht durch LISREL geschätzt wurde. Die Modification Indizes können als Hinweise für Modellverbesserungen herangezogen werden (Abbildung 8).

Abbildung 8: Modification Indices



Die für dieses Modell gefunden Modification Indizes legen eine Modifikation des Meßmodells für SWE (Selbstwirksamkeitserwartungen) nahe: Der Modellfit könnte verbessert werden, wenn die beiden Indikatoren der Prüfungsangst (WOR & EMO) zusätzlich als Indikatoren für die Selbstwirksamkeitserwartungen herangezogen würden. Dies erscheint jedoch theoretisch wenig plausibel.

Zum Schluß wird der globale Modellfit anhand der Fitindizes beurteilt (Tabelle 4). Auch wenn der χ^2-Wert im Verhältnis zu den Freiheitsgraden des Modells etwas zu hoch ausfällt, deuten die übrigen Indizes auf eine gute Modellanpassung hin.

Tabelle 4: Globale Fitindizes

	χ^2	df	p	GFI	NFI	SRMR
Kausalmodell	54	6	.00	.98	.93	.024

MEHRGRUPPENVERGLEICHE

Mit LISREL können Variablenzusammenhänge in unterschiedlichen Stichproben simultan untersucht werden. Diese Option bietet sich z.B. für kulturvergleichende Studien an. Eine mögliche Fragestellung wäre etwa, ob die Korrelation zwischen Berufserfolg und Intelligenz über die Kulturen variiert.

Im folgenden Beispiel soll untersucht werden, ob die Korrelation zwischen Selbstwirksamkeitserwartung und Prüfungsangst für Mädchen und Jungen identisch ist. Die Selbstwirksamkeitserwartung wird mit 7 Items, die Prüfungsangst mit 2 Items operationalisiert. Es werden zwei alternative Modelle aufgestellt. In Modell A ist die Korrelation identisch, in Modell B ist sie es nicht. Eine Entscheidung über die Modelle soll anhand des globalen Fits getroffen werden.

Ein Mehrgruppenvergleich läßt sich ohne großen Aufwand realisieren. Zunächst wird für jede Untersuchungsstichprobe eine Ausgangsmatrix erstellt (hier eine Matrix für Mädchen und eine Matrix für Jungen). Die nachfolgende Syntax realisiert Modell A mit identischer Korrelation zwischen Selbstwirksamkeitserwartung und Prüfungsangst in beiden Stichproben. In der DA(TA)-Zeile wird mit NG=2 die *Number of Groups* festgelegt. LISREL erwartet jetzt zwei Sub-Modelle, die jeweils mit einer DA-Zeile beginnen. Das erste Sub-Modell ist mit MALE betitelt, das zweite mit FEMALE. In der MO-Zeile des Sub-Modells für die Mädchen wird mit LX=SP (SP= same pattern) festgelegt, daß in der LAMBDA-X-Matrix für die Mädchenstichprobe die gleichen Parameter geschätzt werden wie in der LAMBDA-X-Matrix für die Jungenstichprobe: Die Konstrukte sollen in beiden Stichproben mit den gleichen Items operationalisiert werden. Auch die THETA-DELTA-Matrix ist in diesem Sinne spezifiziert. Die PHI-Matrix ist jedoch mit PH = IN(VARIANT) invariant über die Stichproben gesetzt: Es werden die gleichen Parameter wie in der PH-Matrix des 1. Sub-Modells bestimmt, zusätzlich werden die Parameter über beide Stichproben gleichgesetzt. Hiermit wird die Annahme umgesetzt, daß die Korrelation zwischen den latenten Variablen in beiden Stichproben identisch ist. Diese Annahme kann für Modell B sehr einfach wieder aufgehoben werden: Anstelle von PH=IN tritt PH=SP.

```
MALE
DA NG=2 NI=9 NO=1415
LA
schul1_2 schul1_3 schul1_4 schul1_5 schul1_6 schul1_7 schul1_8  tai_e1
tai_w1/
PM SY FI=male.pm
MO NX=9 NK=2
LK
SWE PA/
fr lx 1 1 lx 2 1 lx 3 1 lx 4 1 lx 5 1 lx 6 1 lx 7 1
fr lx 8 2 lx 9 2
OU ML SC

FEMALE
DA NG=2 NI=9 NO=1337
LA
schul1_2 schul1_3 schul1_4 schul1_5 schul1_6 schul1_7 schul1_8  tai_e1
tai_w1/
PM SY FI=female.pm
MO NX=9 NK=2 LX=SP PH=IN TD=SP
LK
SWE PA/
PATH DIAGRAMM
OU ML SC
```

LISREL berechnet einen Modellfit, der sich auf beide Sub-Modelle bezieht. Sollte die Korrelation zwischen Selbstwirksamkeitserwartung und Prüfungsangst für Mädchen und Jungen tatsächlich identisch sein, wird ein guter Fit für Modell A erwartet. Wie der Tabelle 5 zu entnehmen ist, weist Modell A einen befriedigenden Modell fit auf. Durch das Freisetzen der Korrelation wird keine wesentliche Verbesserung erzielt (χ^2-Differenztest), und der RMSEA verschlechtert sich sogar.

Tabelle 5: Modellfit

	χ^2 (df)	RMSEA	GFI
Modell A (invariante Korrelation)	348 (53)	.045	.97
Modell B (freigesetzte Korrelation)	346 (50)	.046	.97

Die Lösung für Modell A ist in Tabelle 6 dargestellt. Die Parameterschätzungen für die PHI-Matrix sind durch PH=IN in beiden Stichproben identisch. Die Korrelation zwischen Selbstwirksamkeitserwartung und Prüfungsangst beträgt demnach für Jungen und Mädchen r = -.31.

Tabelle 6: Standardisierte Lösung für Jungen und Mädchen

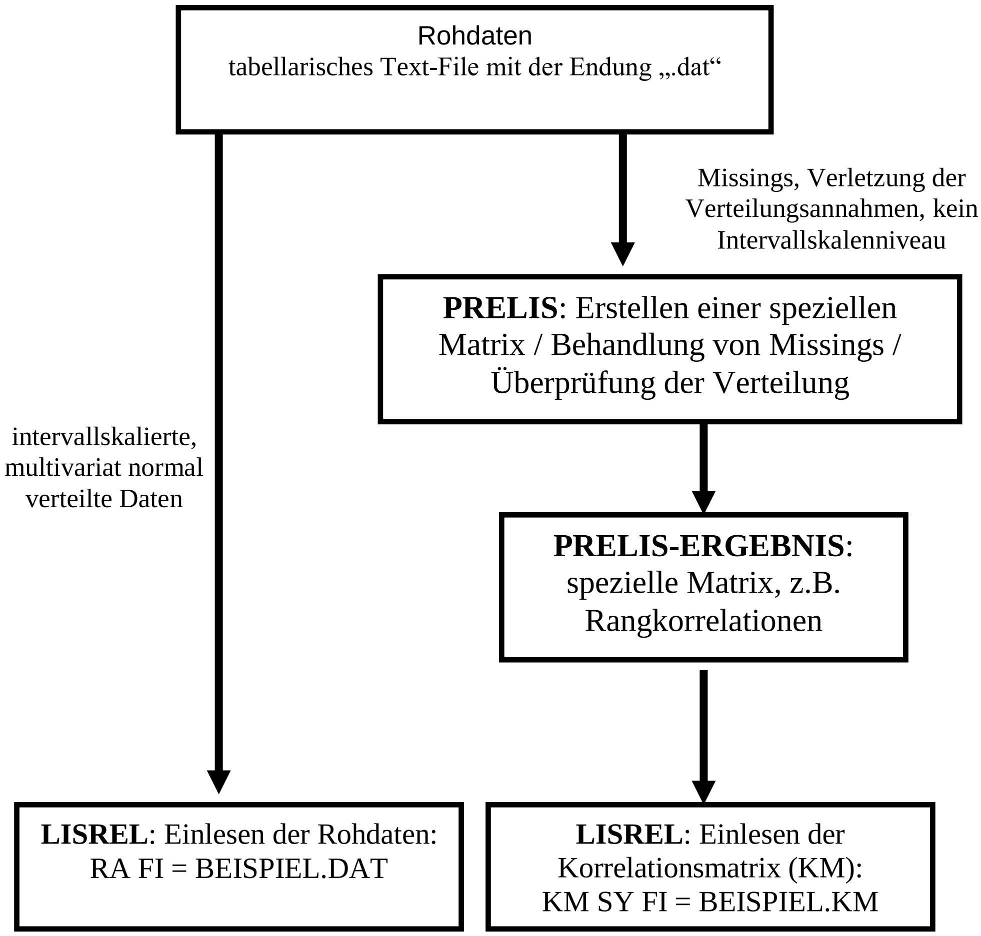
```
MALE
 WITHIN GROUP STANDARDIZED SOLUTION

         LAMBDA-X

                  SWE          PA
             --------    --------
 schul1_2       0.57         - -
 schul1_3       0.63         - -
 schul1_4       0.53         - -
 schul1_5       0.63         - -
 schul1_6      -0.39         - -
 schul1_7       0.47         - -
 schul1_8       0.53         - -
   tai_e1        - -        0.67
   tai_w1        - -        0.89

         PHI

                  SWE          PA
             --------    --------
      SWE       1.00
       PA      -0.31        1.00
FEMALE
 WITHIN GROUP STANDARDIZED SOLUTION

         LAMBDA-X

                  SWE          PA
             --------    --------
 schul1_2       0.61         - -
 schul1_3       0.63         - -
 schul1_4       0.61         - -
 schul1_5       0.71         - -
 schul1_6      -0.44         - -
 schul1_7       0.48         - -
 schul1_8       0.61         - -
   tai_e1        - -        0.75
   tai_w1        - -        0.86

         PHI

                  SWE          PA
             --------    --------
      SWE       1.00
       PA      -0.31        1.00
```

PRELIS

Mit Hilfe von PRELIS kann die Datenaufbereitung vor der LISREL-Analyse optimiert werden. PRELIS ist ein eigenständiges Programm und als Preprozessor zu LISREL gedacht. Neben der Überprüfung von Verteilungseigenschaften und die Erstellung von

Stichprobenmatrizen dient PRELIS dem Umgang mit fehlenden Daten. Ausgang für PRELIS ist eine Textdatei mit Rohwerten (Abbildung 9).

Abbildung 9: Die Funktion von PRELIS



Mit folgender Syntax können die Variablen auf ihre Normalverteilung hin überprüft werden. In der RAW-Zeile wird die Datei mit den Rohdaten spezifiziert. In der CONTINOUS-Zeile werden die Variablen genannt, die als kontinuierlich aufgefaßt werden sollen. In der Output Zeile wird festgelegt, daß eine Kovarianzmatrix erzeugt werden soll:

```
TEST DER VERTEILUNGSEIGENSCHAFTEN
DA NI = 3   MISSING = 99    TREATMENT = PAIRWISE
RAW-DATA = MATRIX.DAT
LABELS
WIRK1 SOZ1 RISK1
CONTINUOUS: WIRK1 SOZ1 RISK1
OUTPUT MATRIX = CMATRIX
```

Ergebnis: Die Variablen wurden auf ihre Normalverteilung hin überprüft. Nur für WIRK1 weichen SKEWNESS und KURTOSIS nicht von einer Normalverteilung ab:

```
TEST OF UNIVARIATE NORMALITY FOR CONTINUOUS VARIABLES
              SKEWNESS          KURTOSIS      SKEWNESS AND KURTOSIS
          Z-SCORE P-VALUE   Z-SCORE P-VALUE   CHI-SQUARE P-VALUE
    WIRK1  -0.742   0.229     0.844   0.199        1.263   0.532
     SOZ1  -2.270   0.012    -2.084   0.019        9.497   0.009
    RISK1  -4.092   0.000     4.369   0.000       35.835   0.000
```

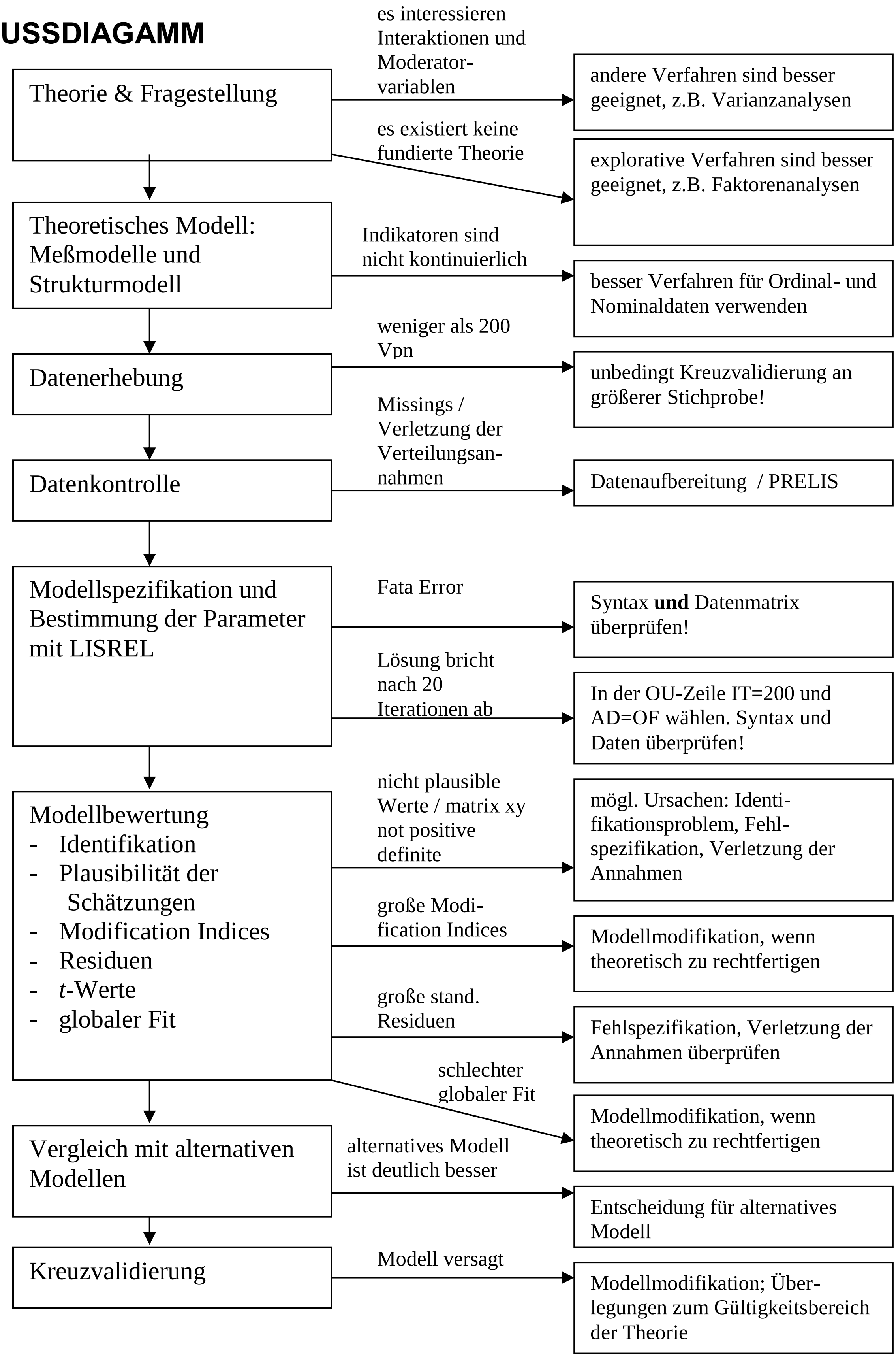
Mit der folgenden Syntax wird aus den Rohdaten eine Kovarianzmatrix erzeugt und in eine Text-Datei geschrieben. Missings (99) werden dabei paarweise ausgeschlossen:

```
TEST DER VERTEILUNGSEIGENSCHAFTEN, ERZEUGEN EINER KOVAR-MATRIX
DA NI = 3   MISSING = 99    TREATMENT = PAIRWISE
RAW-DATA = MATRIX.DAT
LABELS
WIRK1 SOZ1 RISK1
CONTINUOUS: WIRK1 SOZ1 RISK1
OUTPUT MATRIX = CMATRIX FILE CM = MATRIX.CM
```

Ergebnis: Eine Kovarianzmatrix:

```
 TEST DER VERTEILUNGSEIGENSCHAFTEN, ERZEUGEN EINER KOVAR-MATRIX;

         COVARIANCE MATRIX
               WIRK1        SOZ1       RISK1
            --------    --------    --------
    WIRK1      0.160
     SOZ1      0.085       0.209
    RISK1      0.033       0.079       0.469

         MEANS
               WIRK1        SOZ1       RISK1
               2.969       3.122       3.382

         STANDARD DEVIATIONS
               WIRK1        SOZ1       RISK1
               0.400       0.457       0.685
```

FLUSSDIAGAMM



LITERATUR

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (1996). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung (8. Aufl.).* Berlin: Springer-Verlag.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588-606.

Bollen, K. A. & Long, S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 1-9). Newbury Park: Sage Publications.

Boomsma, A. (1987). The robustness of maximum likelihood estimation in structural models. In P. Cuttance & R. Ecob (Eds.), *Structural modeling by example* (pp. 161-187). Cambridge: University Press.

Bortz, J. (1989). *Statistik für Sozialwissenschaftler, 3. Aufl.* Berlin: Springer.

Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation für Sozialwissenschaftler, 2. Aufl.* Berlin: Springer.

Brandmaier, R. & Rietz, C. (Hrsg.). (1993). Methodische Grundlagen und Anwendungen von Strukturgleichungsmodellen (Bd. 2). Mannheim: FRG.

Browne, M. W. & Cudeck, R. (1993). Alternatives ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (S. 136-162). Newbury Park: Sage.

Cuttance, P. (1987). Issues and problems in the application of structural equation models. In P. Cuttance & R. Ecob (Eds.), *Structural modeling by example* (pp. 241-280). Cambridge: University Press.

Cuttance, P. & Ecob, R. (Eds.). (1987). *Structural modeling by example: Applications in educational, and behavioral research*. New York: Cambridge University Press.

Dillon, W., Kumar, A. & Mulani, N. (1987). Offending estimates in covariance structure analysis - comments on the causes and solutions to heywood cases. *Psychological Bulletin*, *101*, 126-35.

Ecob, R. & Cuttance, P. (1987). An overwiev of structural equation modeling. In P. Cuttance & R. Ecob (Eds.), Structural modelling by example (pp. 9-23). Cambridge: University Press.

Ecob, R. (1987). Applications of structural equation modeling to longitudinal data. In P. Cuttance & R. Ecob (Hrsg.), *Structural modeling by example* (S. 138-159). Cambridge: Cambridge University Press.

Engel, U. & Strohe, H. (Hrsg.). (1997). *Hierarchische und dynamische Modellierung. Grundlagen und Anwendungen komplexer Strukturgleichungsmodelle.* Hamburg: Kovac.

Goldberger, A. S. (1973). Structual equation models: An overview. In A. S. Goldberger & O. D. Duncan (Eds.), *Structual equation models in the social sciences*. New York: Seminar Press.

Hair, J. F. (1995). *Multivariate data analysis*. London: Prentice-Hall.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essential and advances*. Baltimore: John Hopkins University Press.

Hox, J. J. & Kreft, T. G. G. (1994). Multilevel analysis methods. *Sociological Methods and Research*, *22*, 283-299.

Hox, J. J. (1993). Mehrebenenanalyse von Strukturgleichungsmodellen. In R. Brandmaier & C. Rietz (Hrsg.), *Methodische Grundlagen und Anwendungen von Strukturgleichungsmodellen* (Bd. 2, S. 36-75). Mannheim: FRG.

Hox, J. J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties.

Jöreskog, K. G. & Sörbom, D. (1988). *PRELIS. A program for multivariate data screening and data summarization. 2ed.* Mooresville, IN: Scientific Software.

James, L. R., Muliak, S. A. & Brett, J. M. (1982). *Causal analysis assumptions, models and data*. Beverly Hills, Calif.: Sage.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36(4)*, 409-426.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-317). Newbury Park: Sage Publications.

Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7: A guide to the program and application*. Chicago: Spss Inc.

Jöreskog, K. G. & Sörbom, D. (1993). *New features in LISREL 8.* Chicago: Scientific Software.

McArdle, J. J. (1996). Current directions in structural factor analysis. *Current Directions in Psychological Science*, *5*, 11-17.

Renkl, A. & Gruber, H. (1995). Erfassung von Veränderung: Wie und wieso? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27*, 173-190.

Revenstorf, D. (1980). *Faktorenanalyse.* Stuttgart: Kohlhammer.

Satow, L. & Schwarzer, R. (1997). Sozial-kognitive Prädiktoren einer gesunden Ernährungsweise: Eine Längsschnittstudie. *Zeitschrift für Gesundheitspsychologie*, *5*, 243-257.

Schnabel, K. U. (1993). Zur Schätzung von Strukturgleichungsmodellen mit hierarchisch geordneten Daten. In R. Brandmaier & C. Rietz (Hrsg.), *Methodische Grundlagen und Anwendungen von Strukturgleichungsmodellen* (Bd. 2, S. 76-93). Mannheim: FRG.

Schnabel, K. U. (1997). Zur Interpretation der vollstandardisierten Lösung bei Längsschnittmodellen. In U. Engel & H. G. Strohe (Hrsg.), *Grundlagen und Anwendungen komplexer Strukturgleichungsmodelle* (S. 109-118). Hamburg: Kovac.

Schneider, W. (1991). Methodische Probleme und Möglichkeiten schulbezogener Längsschnittforschung. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung: Ein Resümee der Längsschnittforschung* (S. 57-82). Stuttgart: Enke.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 10-40). Newbury Park, CA: Sage.