Psychologische Testtheorie

Sitzung 5

Testtheoretische Modelle III

Wahr-oder-Falsch Aussagen zur letzten Sitzung

1. Das parallele Modell und das τ-äquivalente Modell unterscheiden sich nur in der Annahme zur Gleichheit der Varianzen der Fehlervariablen.

2. Die additive Konstante σ_i im essentiell parallelen, essentiell τ-äquivalenten und τ- kongenerischen Modell bezeichnet man als Steigungsparameter.

3. Im τ-äquivalenten Modell ist der Parameter β_i für alle Items i auf 0 fixiert.

4. Ihnen liegt ein räumlicher Intelligenztest vor, für den die folgende Beschreibung gilt: Jedes Item misst die latente Variable gleich gut und gleich messgenau, aber die Itemschwierigkeiten unterscheiden sich. Entsprechend gilt für diesen Test das essentiell parallele Modell.

5. Eine zufällig gezogene Person beantwortet zwei Items eines Empathietests, für den das essentiell τ-äquivalente Modell gilt. Sie antwortet auf das erste Item mit 6 und auf das zweite mit 5. Die Itemparameter sind 1.2 für das erste Item und 1.9 für das zweite Item. Die konkret gezogene Person hat eine Empathie von 3.8.

   a) Der Erwartungswert der Itemantwort einer zufällig gezogenen Person auf Item 1 ist 1.2.

   b) Die Realisation der Zufallsvariable ε_2 ist 1.9.

Überblick

Sitzung	Datum	Thema	Themenblock
1	13.10.25	Einführung	Begriffe, Modellierung von Antwortverhalten durch Zufallsvariablen & mathematische Grundlagen der Testtheorie
2	20.10.25	Wahrscheinlichkeitstheoret. Grundlagen	
3	27.10.25	Testtheoretische Modelle I	Testtheoretische Modelle
4	03.11.25	Testtheoretische Modelle II	
5	10.11.25	Testtheoretische Modelle III	

➛ In der heutigen Vorlesung werden wir uns noch ein letztes testtheoretisches Modell sowie ein paar allgemeinere Aspekte dazu anschauen und wir verschaffen uns einen ersten Überblick über die Gütekriterien psychologischer Tests

2. Testtheoretische Modelle

- Zusammenfassung aller bisherigen Modelle -

Zusammenfassung der bisherigen Modelle I

- Zusammenfassung der jeweils ersten Annahmen der Modelle, d.h. der Annahmen, die sich auf den Zusammenhang zwischen zufälligen latenten Variablen und zufälligen wahren Werten bzw. Itemantworten beziehen:

  - parallel: $X_i = \theta + \varepsilon_i$ da $\tau_i = \theta$
  - essentiell parallel: $X_i = \sigma_i + \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \theta$
  - τ -äquivalent: $X_i = \theta + \varepsilon_i$ da $\tau_i = \theta$
  - essentiell τ -äquivalent: $X_i = \sigma_i + \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \theta$
  - τ -kongenerisch: $X_i = \sigma_i + \beta_i \cdot \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \beta_i \cdot \theta$

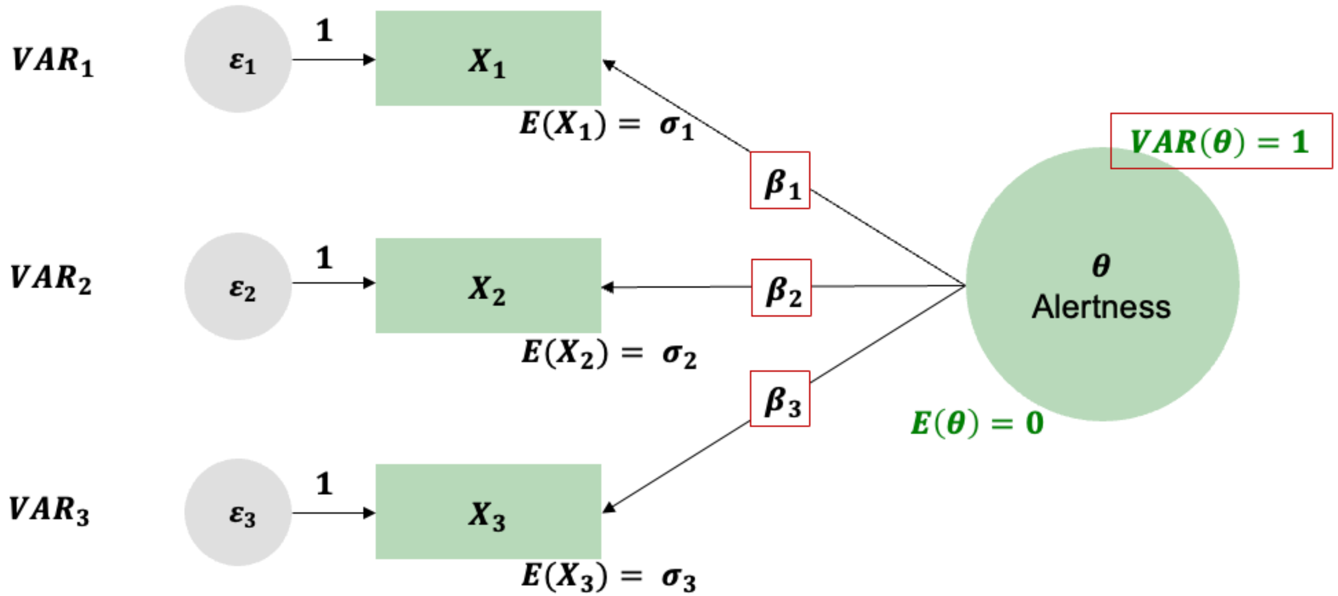
Zusammenfassung der bisherigen Modelle II

Modell	Itemparameter	Steigungs-parameter	Gleiche Fehlervarianzen
parallel	nein	nein	ja
essentiell parallel	ja	nein	ja
τ-äquivalent	nein	nein	nein
essentiell τ-äquivalent	ja	nein	nein
τ-kongenerisch	ja	ja	nein

- Alle vier Modell nehmen zudem $COV(\varepsilon_i, \varepsilon_j) = 0$ für alle Itempaare *i,j* an

Interpretation der Parameter im τ-kongenerischen Modell

- Fragen:
  - Was sagen unterschiedliche Itemparameter über die Items aus?
  - Was sagen unterschiedliche Steigungen über die Items aus?
  - Was sagen unterschiedliche Fehlervarianzen über die Items aus?



2. Testtheoretische Modelle

2.1. Paralleles Modell

2.2. Essentiell Paralleles Modell

2.3. τ-äquivalentes Modell

2.4. Essentiell τ -äquivalentes Modell

2.5. τ -kongenerisches Modell

2.6. Mehrdimensionales τ -kongenerisches Modell

- Alle bisherigen Modelle gingen davon aus, dass nur eine latente Variable das Antwortverhalten der Personen beeinflusst:
  - parallel: $X_i = \theta + \varepsilon_i$ da $\tau_i = \theta$
  - essentiell parallel: $X_i = \sigma_i + \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \theta$
  - τ -äquivalent: $X_i = \theta + \varepsilon_i$ da $\tau_i = \theta$
  - essentiell τ -äquivalent: $X_i = \sigma_i + \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \theta$
  - τ -kongenerisch: $X_i = \sigma_i + \beta_i \cdot \theta + \varepsilon_i$ da $\tau_i = \sigma_i + \beta_i \cdot \theta$
- Modelle dieser Art werden **eindimensionale testtheoretische Modelle** genannt
- Bei einigen Tests ist jedoch anzunehmen, dass mehr als eine latente Variable das Antwortverhalten der Personen erklärt bzw. vorhersagt
- Beispielsweise könnte das Antwortverhalten in einem IQ-Test nicht nur von Intelligenz, sondern auch von Konzentration abhängen
- Modelle mit mehreren latenten Variablen nennen wir **mehrdimensionale testtheoretische Modelle**

- **Eindimensionale** testtheoretische Modelle hatten wir als einfache lineare Regressionsgleichungen mit der latenten Variable als Prädiktor formuliert
- Es liegt also nahe, **mehrdimensionale** testtheoretische Modelle als multiple lineare Regressionsmodelle mit mehreren latenten Variablen als Prädiktoren zu formulieren
- Dies führt uns zum letzten Modell, das wir in diesem Semester besprechen werden: dem **mehrdimensionalen τ-kongenerischen Modell**

Mehrdimensionales τ-kongenerisches Modell: Annahmen

- Im ***q*-dimensionalen** τ-kongenerischen Modell werden folgende Annahmen getroffen:

$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q$ und somit

$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q + \varepsilon_i$ für alle Items i

$COV(\varepsilon_i, \varepsilon_j) = 0$ für alle Itempaare i, j

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme I

$$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q \text{ und somit}$$

$$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q + \varepsilon_i \text{ für alle Items } i$$

- Die erste Annahme bezieht sich auf den Zusammenhang der q zufälligen latenten Variablen $\theta_1, \theta_2, \dots, \theta_q$ mit den zufälligen wahren Werten τ_i der Items des Tests
- In Worten: „Der zufällige wahre Wert auf jedem Item i setzt sich zusammen aus der Summe der jeweils mit einer itemspezifischen Konstante β_{il} gewichteten zufälligen latenten Variablen plus einer itemspezifischen Konstante σ_i"
- Der Index l steht für eine beliebige der q latenten Variablen (so wie i für ein beliebiges der k Items eines Tests steht)
- Die Konstanten σ_i werden wieder **Itemparameter** genannt.
- Die Konstanten β_{il} werden **Steigungsparameter** genannt. Es gibt für jedes Item genau q Steigungsparameter

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme II

$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ldots + \beta_{iq} \cdot \theta_q$ und somit

$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ldots + \beta_{iq} \cdot \theta_q + \varepsilon_i$ für alle Items i

- Wichtige Frage: Woher wissen wir, wie viele latente Variablen die Itemantworten beeinflussen? In anderen Worten: Welchen Wert hat q?
- Bei der Beantwortung dieser Frage gibt es mehrere Vorgehensweisen:
  - Erste Möglichkeit: Wir haben eine Theorie darüber, wie viele latente Variablen hinter einem Test stehen könnten
    - z.B. die Big Five: In diesem Fall würden wir q gleich dem Wert setzen, der sich aus der Theorie ergibt (hier = 5) und dann an einer Stichprobe **überprüfen**, ob das Modell „passt“ (siehe Vorlesung #06)
  - Zweite Möglichkeit: Wir verwenden statistische Methoden um mithilfe von Stichprobendaten **herauszufinden**, wie viele latente Variablen hinter dem Test stehen (siehe Vorlesung #08)

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme III

$$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q \text{ und somit}$$

$$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q + \varepsilon_i \text{ für alle Items } i$$

- Im mehrdimensionalen τ-kongenerischen Modell müssen wir aus Normierungsgründen die Erwartungswerte aller latenten Variablen auf 0 und die Varianzen aller latenten Variablen auf 1 setzen:

$$E(\theta_l) = 0 \text{ für jede latente Variable } l = \ 1, 2, \dots, q$$

$$VAR(\theta_l) = 1 \text{ für jede latente Variable } l = \ 1, 2, \dots, q$$

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme IV

$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q$ und somit

$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ \dots + \beta_{iq} \cdot \theta_q + \varepsilon_i$ für alle Items i

- Zudem müssen wir aus Normierungsgründen (zunächst) die Kovarianzen zwischen den latenten Variablen auf 0 setzen:

$COV(\theta_l, \theta_m) = 0$ für alle latenten Variablen $l \neq m$

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme V

$\tau_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ldots + \beta_{iq} \cdot \theta_q$ und somit

$X_i = \sigma_i + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \ldots + \beta_{iq} \cdot \theta_q + \varepsilon_i$ für alle Items i

$COV(\theta_l, \theta_m) = 0$ für alle latenten Variablen $l \neq m$

- Dies wirkt auf den ersten Blick nicht wie eine einfache Normierung, sondern wie eine (sogar sehr strenge) Modellannahme. Wäre z.B. θ_1 Intelligenz und θ_2 Konzentration, dann würde $COV(\theta_1, \theta_2) = 0$ bedeuten, dass es keinen Zusammenhang zwischen Intelligenz und Konzentration gibt. Dies wäre eine inhaltliche psychologische Aussage, die natürlich falsch sein kann!
- Wir werden jedoch später sehen, dass wir diese Normierungsbedingung nur vorübergehend aus mathematischen Gründen benötigen und später in der Modellierung aufgeben können (Vorlesung #08)

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme VI

- Interpretation der Schwierigkeitsparameter:
  - Wegen der Normierung lassen sich die **Itemparameter** wieder als **Erwartungswerte der Itemantworten** interpretieren:

$$E(X_i) = \sigma_i$$

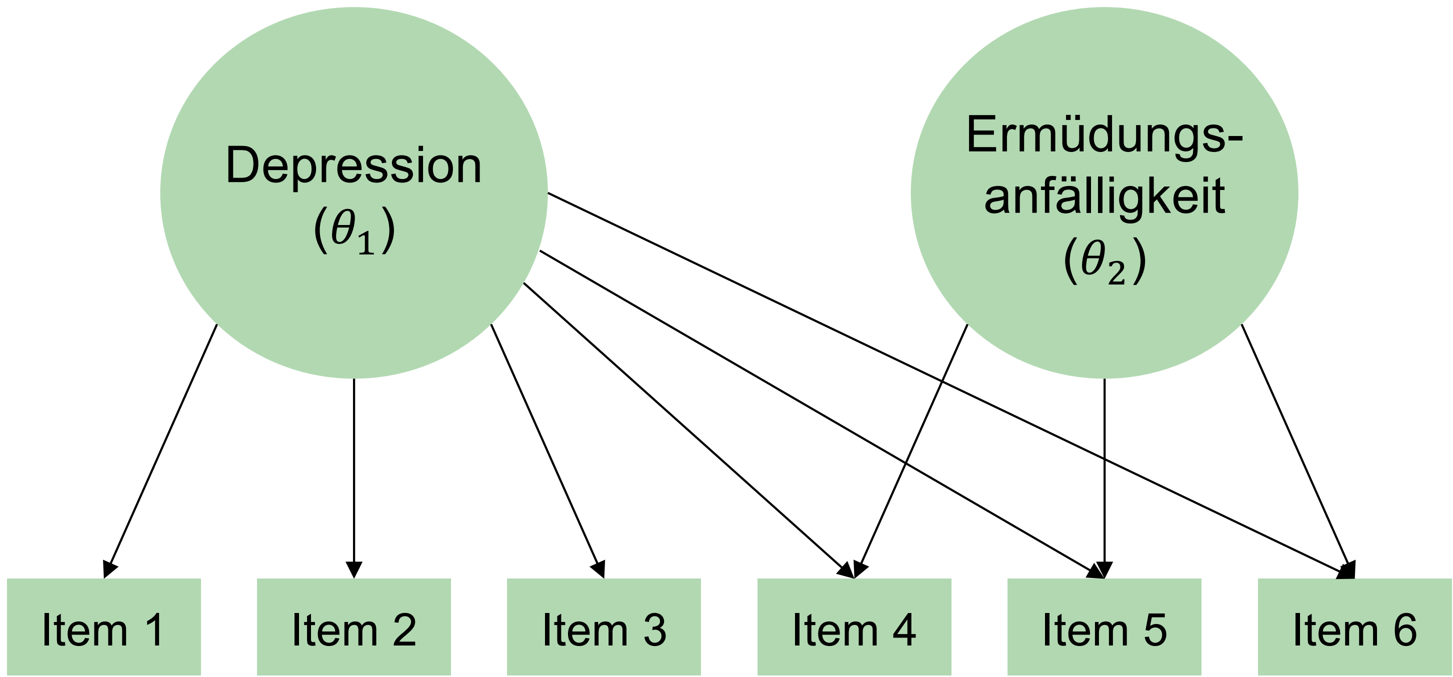
  - Der Beweis funktioniert genau wie in den (eindimensionalen) essentiell parallelen, essentiell τ-äquivalenten und τ-kongenerischen Modellen (siehe Anhang von Sitzung #04)

Mehrdimensionales τ-kongenerisches Modell – Erste Annahme VII

- Interpretation der Steigungsparameter:
  - Der Steigungsparameter β_{il} einer latenten Variable l auf einem Item i entspricht einem **Regressionsgewicht**, das wir schon aus Statistik II kennen
  - Für jedes Item und jede latente Variable hat er somit folgende Interpretation: „Falls sich die latente Variable l um eine Einheit erhöht, erhöht sich die durchschnittliche Itemantwort auf Item i um β_{il} Einheiten, falls alle anderen latenten Variablen des Modells konstant bleiben“
  - Die Höhe von β_{il} gibt jeweils die **Stärke des Zusammenhangs** zwischen der latenten Variable l und der Antwort auf Item i an
  - $\beta_{il} = 0$ bedeutet, dass Item i nicht mit der latenten Variable l zusammenhängt
  - Beispiel: Persönlichkeitstest mit 10 Items, 3-dimensionales Modell, $\beta_{42} = 0.5$: „Falls sich die zweite latente Variable um einen Punkt auf der Ratingskala erhöht, erhöht sich die durchschnittliche Itemantwort auf Item 4 um 0.5 Punkte auf der Ratingskala, falls die erste und dritte latente Variable konstant bleiben“

Beispiele für verschiedene Arten von Mehrdimensionalität I

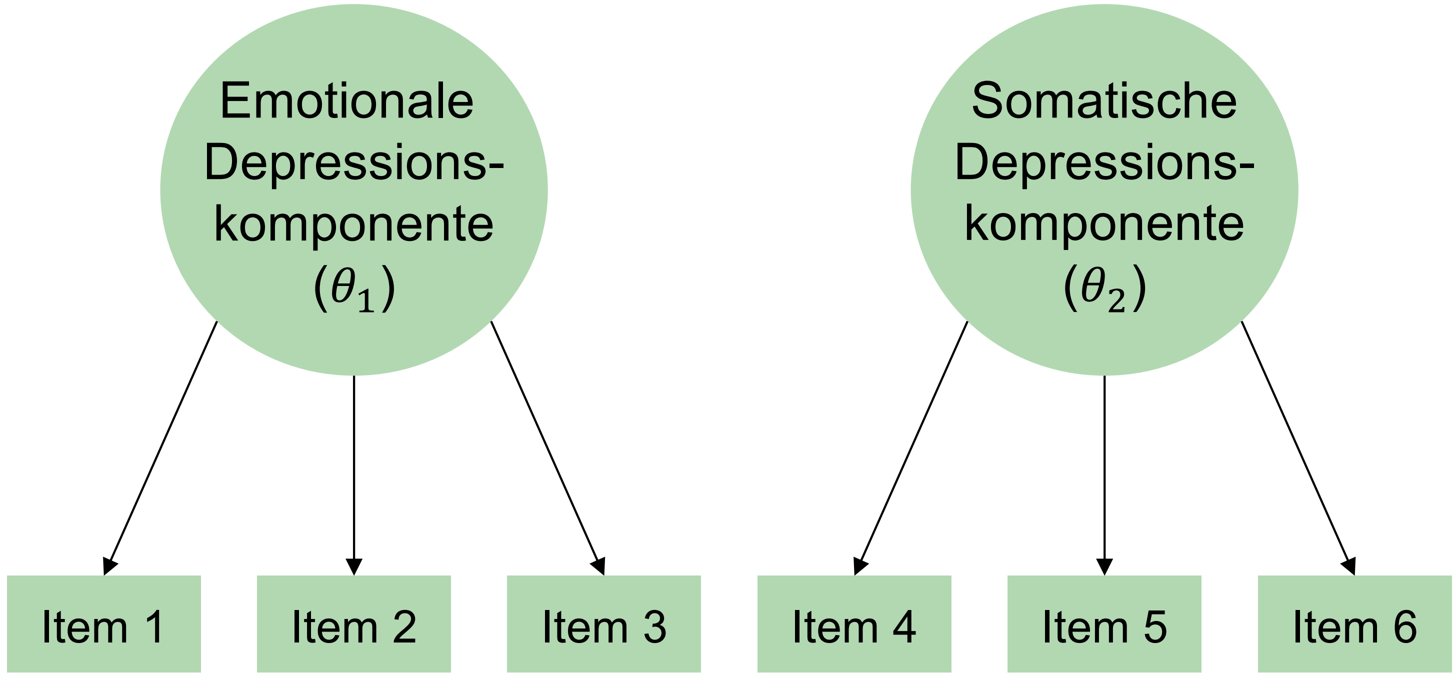
- In dem hier dargestellten Fall werden die ersten drei Items nur durch die latente Variable *Depression* beeinflusst. Für die letzten drei Items spielt neben Depression auch die latente Variable *Ermüdungsanfälligkeit* eine Rolle.



Ein Pfeil in der Grafik bedeutet, dass der Steigungsparameter der jeweiligen latenten Variable <u>nicht gleich 0</u> ist!

Beispiele für verschiedene Arten von Mehrdimensionalität II

- Hier lassen sich die Items des Tests in zwei Gruppen aufteilen, die jeweils nur mit einer der latenten Variablen zusammenhängen. Man kann den Test also in zwei eindimensionale τ-kongenerische Modelle zerlegen.



Ein Pfeil in der Grafik bedeutet, dass der Steigungsparameter der jeweiligen latenten Variable <u>nicht gleich 0</u> ist!

Mehrdimensionales τ-kongenerisches Modell: z-standardisierte Items I

- Im Rahmen der Schätzung des mehrdimensionalen τ-kongenerischen Modells werden häufig **z-standardisierte Itemantworten** betrachtet
- Zur Erinnerung (Statistik I): Jede Zufallsvariable X kann z-standardisiert werden, indem man von ihr ihren Erwartungswert $E(X)$ abzieht und diese Differenz durch die Standardabweichung der Zufallsvariable (Wurzel ihrer Varianz $VAR(X)$) teilt:

$$Z = \frac{X - E(X)}{\sqrt{VAR(X)}}$$

- Aufgrund der Normierung der latenten Variablen entsprechen die Erwartungswerte der Items $E(X_i)$ im mehrdimensionalen τ-kongenerischen Modell jeweils den Itemparametern σ_i , sodass die z-standardisierten Items Z_i folgende Form haben:

$$Z_i = \frac{X_i - E(X_i)}{\sqrt{VAR(X_i)}} = \frac{X_i - \sigma_i}{\sqrt{VAR(X_i)}}$$

- Die z-standardisierten Itemantworten haben dann einen Erwartungswert von $E(Z_i) = 0$ und eine Varianz von $VAR(Z_i) = 1$

Mehrdimensionales τ-kongenerisches Modell: z-standardisierte Items II

- Falls die <u>nicht</u> z-standardisierten Items einem mehrdimensionalen τ-kongenerischen Modell folgen, dann folgen auch die z-standardisierten Items einem mehrdimensionalen τ-kongenerischen Modell, jedoch hat dieses andere Steigungsparameter β_{zil} und Fehlervariablen ε_{zi} und keine Schwierigkeitsparameter mehr (Beweis im Anhang):

$$Z_i = \beta_{zi1} \cdot \theta_1 + \beta_{zi2} \cdot \theta_2 + \cdots + \beta_{ziq} \cdot \theta_q + \varepsilon_{zi} \text{ für alle Items } i$$

- Die **Steigungsparameter $\boldsymbol{\beta_{zil}}$** der z-standardisierten Items werden auch als **Ladungen** bezeichnet und spielen eine wichtige Rolle in der Faktorenanalyse (siehe Sitzungen #08-09)
- Sie haben für jedes Item und jede latente Variable folgende Interpretation: „Falls sich die latente Variable l um eine <u>Standardabweichung</u> erhöht, erhöht sich die durchschnittliche Itemantwort auf Item i um β_{zil} <u>Standardabweichungen</u>, falls alle anderen latenten Variablen des Modells konstant bleiben"

Mehrdimensionales τ-kongenerisches Modell – Fehlervarianzen und zweite Annahme

$COV(\varepsilon_i, \varepsilon_j) = 0$ für alle Itempaare i, j

- Die Varianzen der Fehlervariablen müssen **nicht** für alle Items gleich groß sein
- Die Fehlervariablen ε_i der Items kovariieren nicht (gilt ebenso für die Fehlervariablen ε_{zi} nach der z-Standardisierung der Items, falls diese Annahmen für die Fehlervariablen ε_i erfüllt sind)

2.7 Allgemeine Aspekte testtheoretischer Modelle

Beziehungen zwischen den Modellen

- Die besprochenen Modelle treffen unterschiedlich strenge Annahmen
- Das Modell mit den am wenigsten strengen Annahmen ist das **mehrdimensionale $\boldsymbol{\tau}$-kongenerische Modell**. Alle anderen Modelle sind in dem Sinne strenger, als dass sie Spezialfälle mit zusätzlichen Annahmen sind
  - Zum Beispiel ist das eindimensionale τ-kongenerisches Modell mathematisch gesehen ein „mehrdimensionales" τ-kongenerisches Modell mit der Zusatzannahme $q = 1$
  - Ebenso ist das τ-äquivalente Modell mathematisch gesehen ein essentiell τ-äquivalentes Modell mit Schwierigkeitsparameter gleich 0 für alle Items
  - Letzteres ist wiederum ein eindimensionales τ-kongenerisches Modell mit Steigungsparameter gleich 1 für alle Items
- Wenn ein strengeres Modell gilt, dann gilt immer auch gleichzeitig jedes weniger strenge Modell. Andersrum gilt dies nicht!
- In der Praxis spricht nichts dagegen, gleich das am wenigsten strenge Modell zu verwenden, da die Modelle sich qualitativ nicht unterscheiden

Personenhomogenität I

- Alle bisher besprochenen Modelle nehmen implizit an, dass für alle untersuchten Personen das gleiche Modell gilt. Dies folgt aus den jeweils ersten Annahmen, da die Form der Gleichung für alle Personen identisch ist. Die Personen unterscheiden sich lediglich in ihren Werten auf der latenten Variable und in ihren Itemantworten. Die durch die Gleichungen ausgedrückte Art des Zusammenhangs zwischen den latenten Variablen und den Itemantworten ist für alle Personen identisch.
- Es kann z.B. nicht sein, dass für einen Teil der Population ein Item negativ mit der latenten Variable zusammenhängt (d.h. negativer Steigungsparameter im τ-kongenerischen Modell) und für einen anderen Teil positiv (d.h. positiver Steigungsparameter im τ-kongenerischen Modell)
- Diese implizite Annahme der Modelle wird auch **Personenhomogenität** genannt

Personenhomogenität II

- Personenhomogenität bedeutet inhaltlich unter anderem, dass alle Personen die Items des Tests in gleicher Art und Weise verstehen müssen
- Diese Annahme kann in der Praxis sehr leicht verletzt sein. Beispiel:

Manchmal setze ich Leuten zu oder schmeichle ihnen, damit sie tun was ich will				
○	○	○	○	○
Starke Ablehnung	*Ablehnung*	*Neutral*	*Zustimmung*	*Starke Zustimmung*

- Hier könnte es gut sein, dass verschiedene Personen das Wort „manchmal“ unterschiedlich interpretieren:
  - Die einen verstehen darunter „mehrmals am Tag“, die anderen verstehen darunter „mehrmals im Monat“
  - Für diese beiden Gruppen würden sich eventuell unterschiedliche Modellgleichungen ergeben

Antwortformat der Items I

- Streng genommen sollten die Modelle, die wir bisher kennengelernt haben, nur für **stetige** Itemantworten verwendet werden (z.B. für Reaktionszeit-Items)
- Die meisten psychologischen Tests bestehen jedoch aus Items mit **diskretem** Antwortformat (z.B. Likertskalen):

Ich halte mich für jemanden, der listig und gerissen ist				
1	2	3	4	5
Starke Ablehnung	*Ablehnung*	*Neutral*	*Zustimmung*	*Starke Zustimmung*

- Hier sind nur endlich viele Antworten (1, 2, 3, 4 und 5) möglich. Die Itemantwort auf dieses Item ist somit diskret
- Für solche Items können die Modelle jedoch zumindest approximativ verwendet werden (und das werden sie auch sehr häufig in der Praxis).
  Zur Frage, wie gut diese Approximation funktioniert, gibt es viel Forschung.

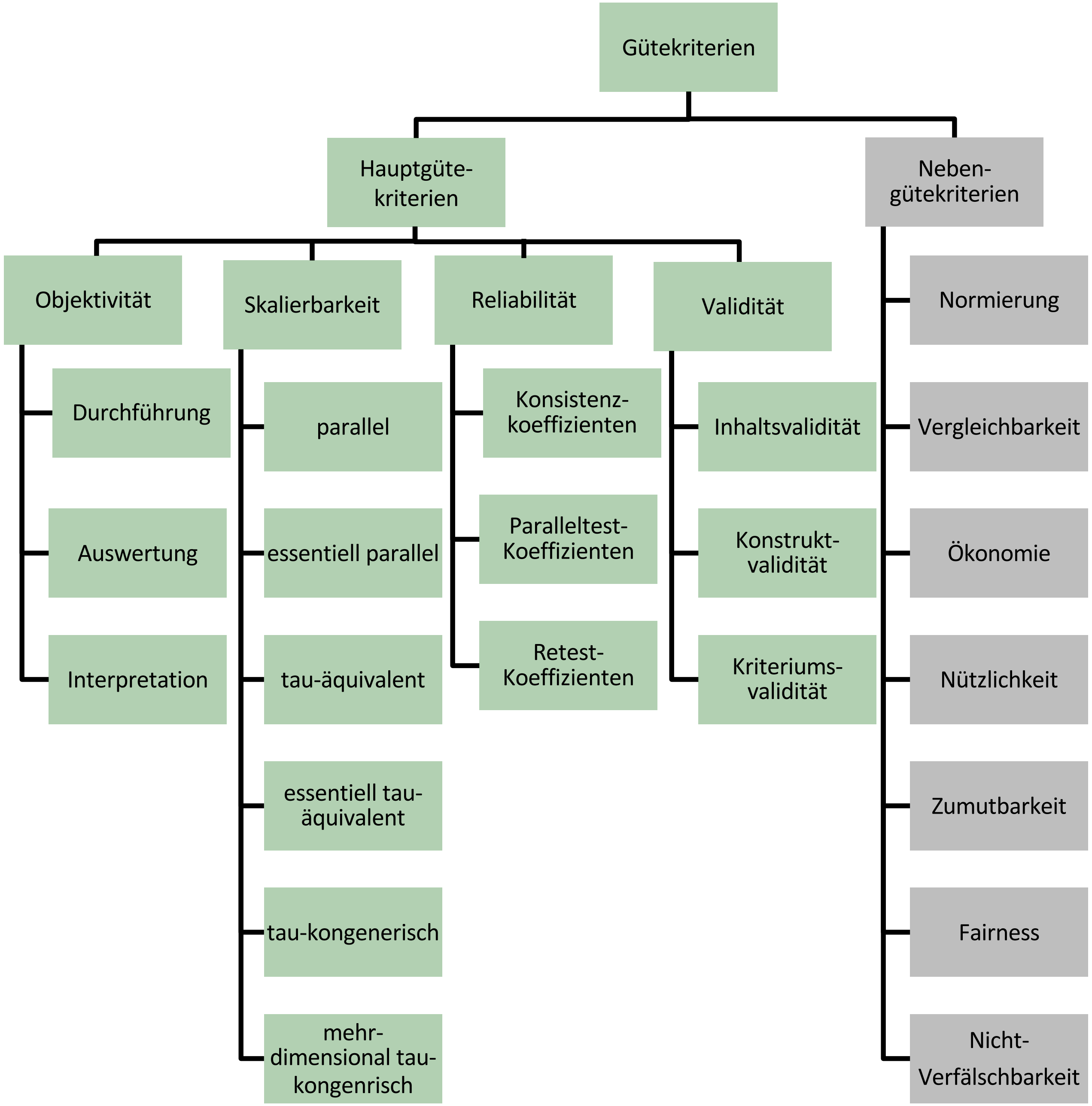
Antwortformat der Items II

- Für einzelne IQ-Test-Items, auf die nur richtig-falsch-Antworten möglich sind, können die hier vorgestellten Modelle nicht direkt angewandt werden. Sie können jedoch für die Summenwerte mehrerer solcher Items verwendet werden:
- Nehmen wir an, wir haben fünf IQ-Items mit ähnlichem Inhalt. Dann wäre die Summe der richtigen Antworten auf diese fünf Items ein neues „Item“, dass die Werte 0, 1, 2, 3, 4 und 5 annehmen könnte. Dieses neue Item wäre dann zwar immer noch nicht stetig, aber immerhin vergleichbar mit dem Item auf der letzten Folie

- Sie haben bisher 6 Modelle kennengelernt, die man der sogenannten „**klassischen Testtheorie**“ zuordnet
- Achtung: Das Wort „klassisch“ soll hier keineswegs andeuten, dass diese Modelle veraltet und nicht mehr relevant sind! Im Gegenteil: Die meisten psychologischen Tests basieren nach wie vor auf diesen Modellen. Eine bessere Bezeichnung wäre zum Beispiel „Testmodelle für kontinuierliche Items“ (siehe vorige Folien).
- Aber: Die klassischen Modelle können nahezu beliebig erweitert werden, z.B. indem man Korrelationen der Fehlervariablen zulässt.
- Und: Zusätzlich wurden über die Jahre Testmodelle für diskrete Itemantworten entwickelt, deren Annahmen vor allem bei dichotomen Intelligenztestitems oder Likertskalen-Items in Fragebögen realistischer sind. Diese Testmodelle, welche häufig unter dem Begriff „Item Response Theory“ oder „Probabilistische Testtheorie“ zusammengefasst werden, lernen Sie im Master kennen

3. Gütekriterien

Übersicht


Gütekriterien
Hauptgüte-kriterien
Neben-gütekriterien
Objektivität
Skalierbarkeit
Reliabilität
Validität
Durchführung
Auswertung
Interpretation
parallel
essentiell parallel
tau-äquivalent
essentiell tau-äquivalent
tau-kongenerisch
mehr-dimensional tau-kongenrisch
Konsistenz-koeffizienten
Paralleltest-Koeffizienten
Retest-Koeffizienten
Inhaltsvalidität
Konstrukt-validität
Kriteriums-validität
Normierung
Vergleichbarkeit
Ökonomie
Nützlichkeit
Zumutbarkeit
Fairness
Nicht-Verfälschbarkeit

Objektivität

Ein psychologischer Test ist dann **objektiv**, wenn die *(1)* Durchführung und *(2)* Auswertung des Tests sowie die *(3)* Interpretation der resultierenden Ergebnisse nicht variieren, auch wenn unterschiedliche Personen den Test zum Einsatz bringen.

- Die Objektivität eines Tests muss primär durch die Rahmenbedingungen der Testanwendung hergestellt werden (siehe Diagnostik II-Vorlesung)
- Es gibt statistische Methoden zur Überprüfung der Objektivität (z.B. Maße der Übereinstimmung mehrerer Testanwender), auf die wir jedoch im Rahmen dieser Veranstaltung nicht eingehen werden (siehe Diagnostik II-Vorlesung)

Skalierung

Ein psychologischer Test gilt als **skalierbar**, wenn die Zuordnung der Messwerte zu den Personen auf der Basis eines empirisch nachgewiesenen testtheoretischen Modells geschieht.

- Wie wir empirisch „nachweisen“ können, dass ein bestimmtes testtheoretisches Modell gilt (also das „richtige“ ist), werden wir in der nächsten Vorlesung (#06) besprechen
- Wie wir auf der Basis eines „nachgewiesenen“ testtheoretischen Modells Personen Werte auf den latenten Variablen zuweisen können, werden wir am Ende des Semesters im Kapitel zur Einzelfalldiagnostik (Vorlesungen #12 und #13) sehen

Reliabilität

Die **Reliabilität** beschreibt die Genauigkeit, mit der ein psychologischer Test ein Merkmal erfasst.

- Was man unter der „Genauigkeit“ eines psychologischen Tests versteht und wie dies mit den testtheoretischen Modellen zusammenhängt, werden wir im Anschluss an die Skalierung (Vorlesungen #09 und #10) besprechen

Validität

Die **Validität** gibt an, ob ein psychologischer Test auch wirklich das misst, was er zu messen beansprucht.

- Verschiedene Aspekte:
  - **Inhaltsvalidität** (Bilden die Items das Konstrukt repräsentativ ab?)
  - **Konstruktvalidität** (Korrelieren Tests, die die gleiche latente Variable messen höher miteinander als Tests, die unterschiedliche latente Variablen messen?)
  - **Kriteriumsvalidität** (Korreliert ein Test mit einem alltagnahen Kriterium? Beispielsweise Berufserfolg gemessen als Einkommen)
- Ob ein psychologischer Test inhaltsvalide ist, ist keine statistische Frage

Das war‘s schon fast für heute…

- *Ausblick:* In der nächsten Vorlesung lernen wir Methoden zur empirischen Überprüfung von testtheoretischen Modellen kennen.

- *Aber zuerst:*
  - **Gibt es offene Fragen zur heutigen Vorlesung?**
  - Zur Vertiefung:
    - Übungsblatt 4 zu allen testtheoretischen Modellen auf Moodle
    - Bühner (2021, Kapitel 4, S. 186-204) für die testtheoretischen Modelle
    - Bühner (2021, Kapitel 8, S. 568-570, 598-599, 600-613) für die Gütekriterien

Anhang:

Mehrdimensionales τ-kongenerisches Modell für z-standardisierte Items

- Bei z-standardisierten Items entfallen im mehrdimensionalen τ-kongenerischen Modell die Schwierigkeitsparameter. Die Steigungsparameter β_{il} werden durch β_{zil} und die Fehlervariablen ε_i durch ε_{zi} ersetzt:

$$Z_i = \beta_{zi1} \cdot \theta_1 + \beta_{zi2} \cdot \theta_2 + \cdots + \beta_{ziq} \cdot \theta_q + \varepsilon_{zi} \text{ für alle Items } i$$

- Beweis:

$$Z_i = \frac{X_i - E(X_i)}{\sqrt{VAR(X_i)}} = \frac{X_i - \sigma_i}{\sqrt{VAR(X_i)}} = \frac{(\cancel{\sigma_i} + \beta_{i1} \cdot \theta_1 + \beta_{i2} \cdot \theta_2 + \cdots + \beta_{iq} \cdot \theta_q + \varepsilon_i) - \cancel{\sigma_i}}{\sqrt{VAR(X_i)}}$$

$$= \underbrace{\frac{\beta_{i1}}{\sqrt{VAR(X_i)}}}_{\boldsymbol{\beta_{zi1}}} \cdot \theta_1 + \underbrace{\frac{\beta_{i2}}{\sqrt{VAR(X_i)}}}_{\boldsymbol{\beta_{zi2}}} \cdot \theta_2 + \cdots + \underbrace{\frac{\beta_{iq}}{\sqrt{VAR(X_i)}}}_{\boldsymbol{\beta_{ziq}}} \cdot \theta_q + \underbrace{\frac{\varepsilon_i}{\sqrt{VAR(X_i)}}}_{\boldsymbol{\varepsilon_{zi}}}$$