2 Messtheoretische und einige weitere methodologische Grundlagen

2.1 Begriffe und Erklärung

2.1.1 Begriffsarten, -intension und -extension

- Zwei Begriffsarten:
  - * logische (und, wenn,...)
  - * empirische (außerlogische):
- Intension (Inhalt): Menge der Merkmale, die für das Vorliegen des Begriffs gegeben sein müssen.
- Extension (Umfang): Menge aller Objekte, die die Intension erfüllen (kann leer sein).

Exkurs: Fuzzy Sets

Kron, T., Winter, L. (2011): Die radikale Unbestimmtheit des Sozialen. In: D. Fischer, W. Bonß, T. Augustin, F. Bader, M. Pichlbauer. D. Vogl (Hg.). Uneindeutigkeit als Herausforderung. Universitätsverlag der Bundeswehr München, Neubiberg, 187-215.[9]

Lemmi, A. Betti, G. (2006): Fuzzy Set Approach to Multidimensional Poverty Measurement. Springer, New York

Smithson, M., Verkuilen, J. (2006): Fuzzy Set Theory: Applications in the Social Sciences. SAGE Publications, New York.

[9] `http://athene-forschung.unibw.de/node?id=89543`, letzter Aufruf 23.10.15

2.1.2 Erklärung, Verifikation, Falsifikation

- Hypothesen und Gesetze

- Theorie: System von Aussagen, das mehrere mit einander in Beziehung stehende, nicht widersprüchliche Hypothesen/Gesetze umfasst

- *Erklärung*: umgangssprachlich vielschichtige Bedeutung des Begriffs

- An wissenschaftliche Erklärung zu stellende Standards

$\rightarrow$ eindeutige Argumentationsstruktur

$\rightarrow$ logisch korrekt

$\rightarrow$ empirisch begründbar

hier nur Reintypen

- Deduktiv-nomologische (D-N-)Erklärung (Nomos: Gesetz), aristotelische Logik, Hempel-Oppenheim Schema

Explanans	Gesetz (Allaussage), Prämisse Randbedingung (Antezedensaussage)		
logische Deduktion			
Explanandum	zu erklärendes Phänomen, Konklusion		

- Erklärung: Explanandum gegeben, wird durch Explanans erklärt

- Vorhersage: Gesetze und Randbedingungen gegeben, Explanandum folgt

Einige Problemfelder

- Richtung des Schlusses!
- inhaltliche Richtigkeit der Konklusion versus logische Richtigkeit des Schlusses.
- Falsifikation

Induktiv-statistische (I-S) Erklärung, probabilistische Erklärung

- Ökologischer Fehlschluss

Wählerwanderungsanalyse							
Wahl '13 → Wahl '09 ↓	CDU/ CSU	SPD	Die Linke	Die Grünen	FDP	Sonstige	Ergebnis 2013↓
CDU/CSU							41,5%
SPD							25,7%
Die Linke							8,6%
Die Grünen							8,4%
FDP							4,8%
Sonstige							11,0%
Ergebnis 2009→	33,8%	23,0%	11,9%	10,7 %	14,6%	6,0%	

2.2 Messen in der Wirtschafts- und Sozialstatistik

2.2.1 Einige Grundbegriffe der repräsentationalen Messtheorie

- Messung im weiteren Sinn: Zuordnung von Zahlen zu Objekten ('Quantifizierung')
- Übliches Ziel: Erhalt von sinnvollen und präzisen Messungen, in welchem Sinn?
- Zur genauen Fassung des Begriffs „Messung“ betrachtet man
  * eine Menge von Objekten mit (mindestens) einer Relation (empirisches Relativ)
  * eine Menge von Zahlen mit (mindestens) einer Relation (numerisches Relativ)
- Eine Messung ist dann eine Abbildung aus dem empirischen in das numerische Relativ

- Jeder Homomorphismus im weiteren Sinn zwischen einem empirischen und einem numerischen Relativ heißt Messung im engeren Sinn (Skala).

  * Ein Homomorphismus ist eine strukturerhaltende Abbildung, bei der also Beziehungen im Sinne der Relation im empirischen Relativ erhalten bleibt.
  * „Parallelität von Sach- und Zahlenlogik“ (Flaskämper)
  * Ein Homomorphismus muss nicht bijektiv sein.

- Liegt zusätzlich ein Isomorphismus vor, so heiße die Skala maximal präzise.

2.2.2 Theoretische Sprache versus Beobachtungssprache, latente Größen und ihre Indikatoren, Operationalisierung

Beispiele für Indikatoren in den Sozialwissenschaften und in der BWL

Beispiele für Wirtschaftsindikatoren, Indikatoren in VWL, empirische Wirtschaftsforschung:

Erste Anforderungen an einen Indikator:

- Plausibilität: theoretischer Zusammenhang zwischen Indikator und latenter Größe
- Datenaktualität: Indikator ist schnell verfügbar, insbesondere für Prognose wichtig
- bei dynamischen Indikatoren zusätzlich
  - statistisch-datentechnische Anforderung: Indikatorreihe soll keine Strukturbrüche aufweisen und verlässlich erhoben worden sein
  - Konformität: die Indikatorreihe spiegelt den Verlauf in der Vergangenheit gut wieder

    später genauere Überlegungen zur Güte von Messungen

Die Festlegung von Korrespondenzregeln, mit deren Hilfe Konstrukte und Indikatoren in Beziehung gesetzt werden, wird insbesondere in der Sozialwissenschaft als *Operationalisierung* bezeichnet.

Vor allem in der amtlichen Statistik wird die Frage der Passung von Konstrukten und Indikatoren als Teilaspekt der *Adäquation* diskutiert. Entsprechende allgemeine Gütekriterien werden später besprochen.

Typischerweise operationalistische Indikatorenbildung

- Direkte *Definition* des Begriffs über Messvorschrift (operationale Definition von Begriffen)
- vgl. physikalische Einheiten: Ein Ampere ist die Stärke eines zeitlich unveränderlichen elektrischen Stromes, der, durch zwei im Vakuum parallel im Abstand 1 Meter voneinander angeordnete, geradlinige, unendlich lange Leiter von vernachlässigbar kleinem, kreisförmigem Querschnitt fließend, zwischen diesen Leitern pro 1 Meter Leiterlänge die Kraft $2 \cdot 10^{-7}$ Newton hervorrufen würde.

2.2.3 Einige weitere Problemfehler

- Reifizierungsproblem:

- Das $Basissatzproblem$:

  „So ist die empirische Basis der objektiven Wissenschaft nichts „Absolutes“; die Wissenschaft baut nicht auf Felsengrund. Es ist eher Sumpfland. . . “ Popper (1976; p. 75) zitiert nach Schnell, Hill, Esser (2011; p. 82)

- multiple Indikatoren

2.2.4 Indizes

Beispiele:

- Human Development Index[10]
- Happy Planet Index [11]
- Times Higher Education Index [12]
- Welthungerindex [13]

Zusammenfassung mehrerer Indikatoren eines Konstrukts zu einer Kennzahl.

[10] http://hdr.undp.org/en/statistics/hdi/, aufgerufen am 15.10.15
[11] http://www.happyplanetindex.org/, aufgerufen am 15.10.15
[12] https://www.timeshighereducation.com/world-university-rankings,aufgerufen am 15.10.15
[13] http://www.welthungerhilfe.de/welthungerindex2015.html, aufgerufen am 23.10.15

- Oft sind Konstrukte mehrdimensional
  - Werden die Dimensionen jeweils durch einen Indikator erfasst?
  - Welche Dimensionen fließen ein?
  - Wie werden sie kombiniert?

- Jeder Ausprägungskombination wird ein Wert zugeordnet. Typische Zuordnungsregeln für metrische Variablen:

  * häufig einfach Verhältniszahlen (vgl. Kap. 3.1)
  * häufig nur Normierung, z.B. Anzahl Kinder im gebärfähigen Alter
  * additiver Index
  * multiplikativer Index
  * gewichteter Index $\sum_{j=1}^{k} g_j \cdot C^{(j)}$, wobei $C^{(j)}$ mit $j = 1, \ldots, k$ die Items bzw. Indikatoren widerspiegeln und g_j die (festgesetzten) Gewichte darstellen
  * auch Preisindizes als gewichtete Indizes
  * häufig einzelne Indikatoren vorher standardisiert („Z-Scores“)
  * Diese Verknüpfungen durch Rechenoperationen als solche setzen ein metrisches Skaleniveau voraus. Daher ist die Verwendung von Indizes in vielen Situationen kritisch zu hinterfragen.

- weiterführend: Skalierungsverfahren, siehe Kap. 2.5

2.3 Das Grundmodell der klassischen Testtheorie (KTT)

Wie kann man die Messproblematik statistisch fassen?
Schnittpunkt zwischen üblicher empirischer Methodik und fortgeschrittenerer statistischer Methodik.

- Man könnte ähnlich wie beim Inferenzfehler vorgehen und sagen: Wenn der Messfehler schon prinzipiell unvermeidlich ist, so sollte man ihn wenigstens statistisch kontrollieren/ mitberücksichtigen.

- Betrachtet wird hier ein Modell für die Messung einer metrischen, potentiell mehrdimensionalen, latenten Größe.
  Notation: Für jeden Vektor $a \in \mathbb{R}^r$ bezeichne $a[\ell]$ die ℓ-te Komponente, also $a = (a[1], a[2], \ldots, a[r])^T$.

- Verallgemeinerungen bezüglich der Messituation bzw. der expliziten Mitberücksichtigung des Messprozess in statistischen Modellen werden später nochmals aufgegriffen (siehe die Abschnitte „Skalierungsverfahren“ (Kap. 2.5) und „Fehler-in-den-Variablen-Modellen“ (in Kap. 7)

2.3.1 Im Kontext einer einzelnen Messung

$$\underset{\substack{\text{gemessener}\\ \text{Wert}}}{\underset{\uparrow}{C}} = \underset{\substack{\text{wahrer}\\ \text{Wert}}}{\underset{\uparrow}{\Gamma}} + \underset{\text{Messfehler}}{\underset{\uparrow}{\Delta}}$$

Da sowohl Γ als auch Δ latent sind, ist diese Gleichung $immer$ erfüllt, man setze einfach $\Delta := T - \Gamma$. Entscheidend sind daher, die Zusatzannahmen, die man unterlegt. Diese sind in konkreten Anwendung unbedingt kritisch zu hinterfragen.

Was bedeutet hier eigentlich "Modell"?

Def. 2.1.

Sei $(\mathbf{C}, \mathbf{\Gamma}, \mathbf{\Delta})$ ein Tripel von metrischen mehrdimensionalen Merkmalen und C, Γ, Δ die daraus durch reine Zufallsauswahl eines Elements gewonnenen dreidimensionale Zufallsvektoren.

1. $(\mathbf{C}, \mathbf{\Gamma}, \mathbf{\Delta})$ und (C, Γ, Δ) mit

$$C = \Gamma + \Delta \tag{2.1}$$

   soll als dem *unzentrierten Grundmodell der klassischen Testtheorie* gehorchend bezeichnet werden, wenn gilt:

(a) $\mathbb{E}(\Delta)$ und $Cov(\Delta)$ existieren.[14]

(b) Γ und Δ sind komponentenweise unabhängig: Für alle $\ell_1 = 1 \ldots, r$ und $\ell_2 = 1 \ldots, r$ gelte $\Gamma[\ell_1]$ und $\Delta[\ell_2]$ sind unabhängig

(c) Messfehler bei verschiedenen Komponenten sind voneinander unabhängig: $\Delta[1], \Delta[2], \ldots, \Delta[r]$ sind stochastisch unabhängig

2. Gilt zusätzlich, dass kein systematischer Messfehler vorliegt, d.h.

$$\mathbb{E}(\Delta) = 0,$$

so sei vom *(zentrierten) Grundmodell der klassischen Testtheorie* gesprochen.

[14] Diese Bedingung wird meist implizit vorausgesetzt.

Bem. 2.2.

Werden einzelne Komponenten $\Gamma[\tilde{\ell}]$ von $\vec{\Gamma}$ messfehlerfrei erhoben, so wird einfach $\Delta[\tilde{\ell}] = 0$ gesetzt. Man kann das Modell erweitern, wenn man vorausgesetzt, dass alle nicht metrischen Variablen exakt gemessen werden, der Ansatz (Nr. 2) passt also auch für sie.

- Man müsste eigentlich noch eine $starke$ und eine $schwache$ Form unterscheiden. Die hier gegebene Formulierung entspräche dann der starken Form; bei der schwachen Form wird nur jeweils Unkorreliertheit statt Unabhängigkeit gefordert.

- Vor allem in der Psychometrie wird allerdings häufig implizit zusätzlich vorausgesetzt, dass $(C, \Gamma, \Delta)^T$ multivariat normalverteilt ist, wodurch starke und schwache Form zusammenfallen.

Def. 2.3.

In der in Definition 2.1 eingeführten Situationen heiße dann jeweils C *Messung des Konstrukts* Γ (mit *Messfehler* Δ) im Sinne des unzentrierten bzw. zentrierten Grundmodells der klassischen Testtheorie.

Erste Konsequenzen

2.3.2 Individuelle Messungen, mehrere Personen

Jetzt Stichprobe von Umfang n

Def. 2.4.

Die in Definition 2.1 eingeführten Bezeichnungen seien auf $i.i.d$ Stichproben $(C_i, \Gamma_i, \Delta_i), \quad i = 1, \cdots, n$, des Tripels $(\mathbf{C}, \mathbf{\Gamma}, \mathbf{\Delta})$ von Merkmalen ausgedehnt, wenn $(C_i, \Gamma_i, \Delta_i)$ für alle i die entsprechenden Bedingungen erfüllt.

2.3.3 Wiederholungsmessungen

Ermittle individuelle Ausprägungen einer latenten Eigenschaft mit Hilfe einer 'Batterie' von Fragen ('Items')

Def. 2.5.

In analoger Weise seien die Bezeichnungen auch auf Wiederholungsmessungen ausgedehnt, wenn zusätzlich gilt

a) im Fall von $(C^{(j)}, \Gamma, \Delta^{(j)}), \quad j = 1, \cdots, p$:

  a1) Für jedes j erfüllt das Tripel $(C^{(j)}, \Gamma, \Delta^{(j)})$ die entsprechenden Bedingungen.

  a2) $(\Delta^{(1)}, \cdots, \Delta^{(j)}, \cdots, \Delta^{(p)})$ sind identisch verteilt und $(\Gamma, \Delta^{(1)}, \cdots, \Delta^{(j)}, \cdots, \Delta^{(p)})$ sind gemeinsam unabhängig.

b) im Fall von $(C_i^{(j)}, \Gamma_i, \Delta_i^{(j)}), \quad i = 1, \cdots, n, \quad j = 1, \cdots, p$:

  b1) Für jedes feste $i, \quad i = 1, \cdots, n$, genügen die Tripel $(C_i^{(j)}, \Gamma_i, \Delta_i^{(j)}), \quad j = 1, \cdots, p$ dem entsprechenden Grundmodell im Sinne von Teil a),

  b2) Für jede beliebige Auswahl $(j_1, \cdots, j_i, \cdots, j_n)$ genügen die Tripel $(C_i^{(j_i)}, \Gamma_i, \Delta_i^{(j_i)}), \quad i = 1, \cdots, n$, dem entsprechenden Grundmodell im Sinne von Def. 2.1.

d) Einige weitere Konsequenzen

Bem. 2.6.

Man kann zeigen: Gehorcht $(C_i^{(j)}, \Gamma_i, \Delta_i^{(j)}), \quad i = 1, \cdots, n;\ j = 1, \cdots, p$ dem unzentrierten (bzw. zentrierten) Grundmodell der klassischen Testtheorie (jeweils in der starken Form):

a) Für jedes i gilt:

$$\overline{C_i} := \frac{1}{p} \sum_{i=1}^{p} C_i^{(j)}$$

genügt dem unzentrierten (bzw. zentrierten) Grundmodell der klassischen Testtheorie.

b) Es gilt die sogenannte *lokale stochastische Unabhängigkeit*:
Für jedes i sind für jedes $j_1 \neq j_2$ die Messungen $C_i^{(j_1)}$ und $C_i^{(j_2)}$ bedingt unabhängig gegeben Γ_i, d.h. es gibt keine über Γ_i hinaus gehende Abhängigkeit zwischen $C_i^{(j_1)}$ und $C_i^{(j_2)}$

2.4 Adäquation; Gütekriterien: Objektivität, Reliabilität, Validität

Eine Messung im schwachen Sinn kann zunächst relativ willkürlich und inhaltlich sinnlos sein.

Bei der Frage nach der Adäquation einr Formalisierung ist die Frage nach der Güte der Messungen ein zentraler Punkt. Bei der Beurteilung unterscheidet man typischerweise drei aufeinander aufbauende Aspekte:

i) Objektivität:

Grad der Unabhängigkeit der Messung von Einflüssen außerhalb untersuchten Einheit

* Durchführungsobjektivität:

* Auswertungsobjektivität:

* Interpretationsobjektivität:

ii) Reliabilität (Zuverlässigkeit)
In welchem Ausmaß führt eine wiederholte Messung zu demselben Ergebnis?

iii) Validität (Gültigkeit)
Grad der Genauigkeit, mit der ein Verfahren oder eine Messung das misst, was es messen soll

Einige Regeln zur Formulierung von Items

- eindimensionale Items
- Frage nach gegenwärtigem Zustand
- keine Suggestion durch Tatsachenbeschreibungen,
- Items sollten den Wertebereich, in dem die Befragten liegen, abdecken
- einfache, klare, kurze, verständliche Struktur, keine Mehrdeutigkeit, keine doppelte Verneinung
- insbesondere natürlich widerspruchsfrei.
- Antwortkategorien erschöpfend

Es sei im Folgenden festgesetzt, dass Γ eindimensional sei:

2.4.1 Beurteilung der Reliabilität

Def. 2.7. *Maßzahl für die Reliabilität (theoretische Reliabilität):*

Gegeben sei eine Messung C des eindimensionalen Konstrukts Γ (mit Messfehler Δ) im Sinne des unzentrierten Grundmodells der klassischen Testtheorie mit $Var(C) > 0$. Dann heißt

$$Rel(\Gamma, C) := \frac{Var(\Gamma)}{Var(C)}$$

$theoretische\ Reliabilität$ von C für Γ.

Satz 2.8.

In der Situation von Def. 2.7 gilt für die Korrelation $\varrho(\Gamma, C)$ von Γ und C:

$$\varrho(\Gamma, C) = \sqrt{Rel(\Gamma, C)}$$

Bem. 2.9.

Alle hier gegebenen Argumentationen stehen und fallen mit der Unkorreliertheit von Γ und Δ !!

„Schätzung“ der Reliabilität

Satz 2.10. [Hauptsatz der Reliabilitätsschätzung] *Seien* $C^{(1)}, C^{(2)}$ *zwei Wiederholungsmessungen für* Γ*, die dem unzentrierten Grundmodell der klassischen Testtheorie im Sinne von Definition 2.5 genügen. Dann gilt für den Korrelationskoeffizienten* $\rho(C^{(1)}, C^{(2)})$

$$\rho(C^{(1)}, C^{(2)}) = Rel(\Gamma, C^{(1)}) = Rel(\Gamma, C^{(2)}).$$

Satz 2.11. *(Spearman-Brown Formel)*

Seien $C^{(1)}, C^{(2)}, \cdots, C^{(q)}$ Wiederholungsmessungen, die dem unzentrierte Grundmodell der klassischen Testtheorie im Sinne von Definition 2.5 genügen. Dann gilt für

$$\overline{C} := \frac{1}{p} \sum_{j=1}^{p} C^{(j)}$$

i) $\overline{C}$ ist eine Messung des Konstrukts Γ im Sinne des unzentrierten Grundmodells der Klassischen Testtheorie.

ii) Es gilt für jedes $j = 1, \cdots, p$

$$Rel(\Gamma, \overline{C}) = \frac{p \cdot Rel(\Gamma, C^{(j)})}{1 + (p-1) Rel(\Gamma, C^{(j)})}$$

Beweis:

Praktische Umsetzung

- $Validierungsdaten$:

- Test-Retest-Methode:

- Paralleltestmethode:

- Split-Half-Methoden zur Erzeugung paralleler Tests:

- Cronbachs Alpha (berücksichtigt die verschiedenen Aufteilungsmöglichkeiten bei p Messungen).
  Normierte Form:
  $$\alpha := \frac{p \cdot \bar{\varrho}}{1 + \bar{\varrho}(p-1)}$$
  mit $\bar{\varrho}$ als durchschnittlicher Korrelationskoeffizient zwischen zwei Messungen

Bsp. 2.12. *Man betrachte die in Abb. 1 gegebene Situation und bestimme Cronbachs* α

Reliabilitätsstatistiken

Cronbachs Alpha	Cronbachs Alpha für standardisierte Items	Anzahl der Items
,799	,801	4

Inter-Item-Korrelationsmatrix

	FRAU, LIEBER MANN BEI D. KARRIERE HELFEN?	FRAU, NICHT ARBEITEN BEI KLEINKIND?	FRAU, ZU HAUSE BLEIBEN+KINDER VERSORGEN?	FRAU, NACH HEIRAT ARBEITSPL. FREIMACHEN?
FRAU, LIEBER MANN BEI D.KARRIERE HELFEN?	1,000	,385	,547	,457
FRAU, NICHT ARBEITEN BEI KLEINKIND?	,385	1,000	,583	,446
FRAU, ZU HAUSE BLEIBEN+KINDER VERSORGEN?	,547	,583	1,000	,587
FRAU, NACH HEIRAT ARBEITSPL. FREIMACHEN?	,457	,446	,587	1,000

Abbildung 1: Beispiel aus dem ALLBUS 2008, bereitgestellt von GESIS, Leibniz-Institut für Sozialwissenschaften.

2.4.2 Beurteilung der Validität

Def. 2.13. *(Theoretische Validität:)*

Sei C eine Messung für Γ, die dem zentrierten Grundmodell der Klassischen Testtheorie genügt. Dann heißt $\rho(C, \Gamma)$ theorethische Validität.
Im Lichte von Satz 2.8 entspricht also die Validitätsmessung der Reliabilitätsmessung, wobei aber zusätzlich gefordert wird,

$$\mathbb{E}(\Delta) = 0,$$

also die zentrierte Variante vorliegt.

- „Axiomatisierung“

- $Inhaltsvalidität$: Alle Dimensionen des Konstrukts erfasst – und nur diese!

Goodhart's Law:

Praktische Beurteilung der Validität:

- *Expertenbefragung*

- *Kriteriumsvalidität*: Hoher Zusammenhang zwischen Messwerten und einem anderen gemessenem Kriterium („externes Kriterium“)
  - Prädiktive Validität:
  - Konkurrente Validität:

- *Konstruktvalidität:*
  - Kriterien:
    Konvergenz: *Diskriminanz*:

* *Multitrait-Multi-Method-Matrix*

2.4.3 Mehrere Merkmale, Korrelations- und Regressionsanalyse

Satz 2.14.

Seien C und T Messungen für Γ bzw. Θ mit Messfehler Δ bzw. Ψ im Sinne des Grundmodells der klassischen Testtheorie und Δ und ψ stochastisch unabhängig. Dann gilt:

$$\varrho(\Gamma, \Theta) = \frac{\varrho(C, T)}{\sqrt{\mathsf{Rel}(\Gamma, C)} \cdot \sqrt{\mathsf{Rel}(\Theta, T)}} \tag{2.2}$$

Beweis:

Bem. 2.15.

Die rechte Seite wird oft als Masszahl der Kriteriumsvalidität verwendet. Sind die Voraussetzungen des Satzes erfüllt, so lässt sich folgern:

1. Unter Messfehlern ist $\text{Rel}(\Gamma, C) < 1$ und/oder $\text{Rel}(\Theta, T) < 1$; also zeigen sich die wahren Zusammenhänge *abgeschwächt* im zugehörigen Indikator (engl. attenuation).

2. $\sqrt{\text{Rel}(\Gamma, C)} \cdot \sqrt{\text{Rel}(\Theta, T)}$ kann als Korrekturfaktor gesehen werden, um von $\varrho(C, T)$ auf $\varrho(\Gamma, \Theta)$ schliessen zu können („correction for attenuation“).

3. Seien $\beta(C, \Theta)$ bzw. $\beta(C, T)$ die “theoretischen Regressionskoeffizienten” einer klassischen linearen Regression mit Absolutglied, Γ bzw. C als abhängige Variable und Θ bzw. T als Kovariable. Dann gilt:

$$\beta(C, T) = \text{Rel}(\Theta, T) \cdot \beta(\Gamma, \Theta) \tag{2.3}$$

Beweis:

4. Die Herleitungen benutzen explizit, dass

$$Cov(\Gamma, \Psi) = Cov(\Delta, \Theta) = Cov(\Delta, \Psi) = 0\,.$$

   Bestehen Abhängigkeiten, so gilt (2.2) nicht mehr. Die entsprechenden Korrekturfaktoren können grösser oder kleiner sein.

5. Die Aussagen gelten gleichermaßen für das unzentrierte Grundmodell der Klassischen Testtheorie, denn die Bedingungen $\mathbb{E}(\Delta) = 0$ und $\mathbb{E}(\Psi) = 0$ werden in der Herleitung nicht benötigt: Valide Messinstrumente und nicht valide Messinstrumente mit derselben Reliabilität liefern dieselbe Korrelation.