

Forschungsmethoden

VORLESUNG WS 2017/2018

SOPHIE LUKES



Überblick

Letzte Woche:

- Stichprobenziehung und Stichprobeneffekte

Heute:

- Gütekriterien I

Rückblick

- Population und Stichprobe
- verschiedene Arten der Stichprobenziehung
- Kontrollgruppen der klinischen Forschung

Heute: Gütekriterien

Warum Gütekriterien?

- Gütekriterien geben einen Maßstab für wissenschaftliches Arbeiten vor
- zur Bewertung von Verfahren und Messinstrumenten

Gütekriterien



Gütekriterien der qualitativen Forschung

- kein einheitlicher Standard wie im quantitativen Ansatz
- verschiedene Kriterienkataloge

Vier Gütekriterien für qualitative Forschung nach Lincoln & Guba (1985)

- Vertrauenswürdigkeit
- Übertragbarkeit
- Zuverlässigkeit
- Bestätigbarkeit

Glaubwürdigkeit

Vier Gütekriterien für qualitative Forschung nach Lincoln & Guba (1985)

- **Vertrauenswürdigkeit**
 - Bezogen auf Ergebnisse und Interpretationen auf Datenbasis
 - Zu erreichen über:
 - Umfassende Datenerhebung
 - Peer Debriefing
 - Reflexion aufgrund von Gegenbeispielen
 - Kommunikative Validierung

Vier Gütekriterien für qualitative Forschung nach Lincoln & Guba (1985)

- **Übertragbarkeit**

- Übertragung der Schlussfolgerungen auf andere Kontexte möglich
- Zu erreichen über:
 - Genaue Kontext- und Personenbeschreibungen zur Einschätzung der Möglichkeit der Übertragbarkeit

Vier Gütekriterien für qualitative Forschung nach Lincoln & Guba (1985)

- **Zuverlässigkeit**

- Forschungsprozess nachvollziehbar gestalten und durchführen
- zu erreichen über:
 - Darstellung gegenüber Experten
 - Prüfung der mit einer Methode gewonnenen Daten durch Verwendung anderer Methode

Vier Gütekriterien für qualitative Forschung nach Lincoln & Guba (1985)

- **Bestätigbarkeit**

- keine Vorbestimmung der Ergebnisse durch Interessen oder Sichtweisen der Forschenden
- Zu erreichen über:
 - Darstellung der Daten und des Forschungsprozesses, zum Beispiel auch durch Dokumente zur Selbstreflexion

Gütekriterien in der quantitativen Forschung

Hauptgütekriterien

Nebengütekriterien

Hauptgütekriterien

- Objektivität
- Reliabilität
- Validität

Nebengütekriterien

- Skalierung
- Normierung
- Testökonomie
- Nützlichkeit
- Zumutbarkeit
- Unverfälschbarkeit
- Fairness

Heute

- **Objektivität**
- **Reliabilität**
- **Validität**

Objektivität

= intersubjektive Vergleichbarkeit

- Testergebnis soll unabhängig vom Testanwender sein

Objektivität

- **Durchführungsobjektivität**
- Auswertungsobjektivität
- Interpretationsobjektivität

Objektivität

- Durchführungsobjektivität
- **Auswertungsobjektivität**
- Interpretationsobjektivität

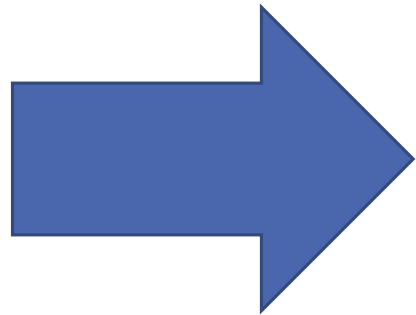
Objektivität

- Durchführungsobjektivität
- Auswertungsobjektivität
- **Interpretationsobjektivität**

Wie kann man die Objektivität maximieren?

- Standardisierung der Testsituation
- Versuchsleiterschulung
- genaue Instruktion
- genaue Richtlinien zur Auswertung und Interpretation
- Normierung

Reliabilität



Beruht auf der Klassischen Testtheorie (KTT)

Klassische Testtheorie

Beispiel 1: Messung des Gewichtes eines Buches mit einer Waage

- Gewicht bleibt gleich/schwankt nur gering, Waage ist reliabel

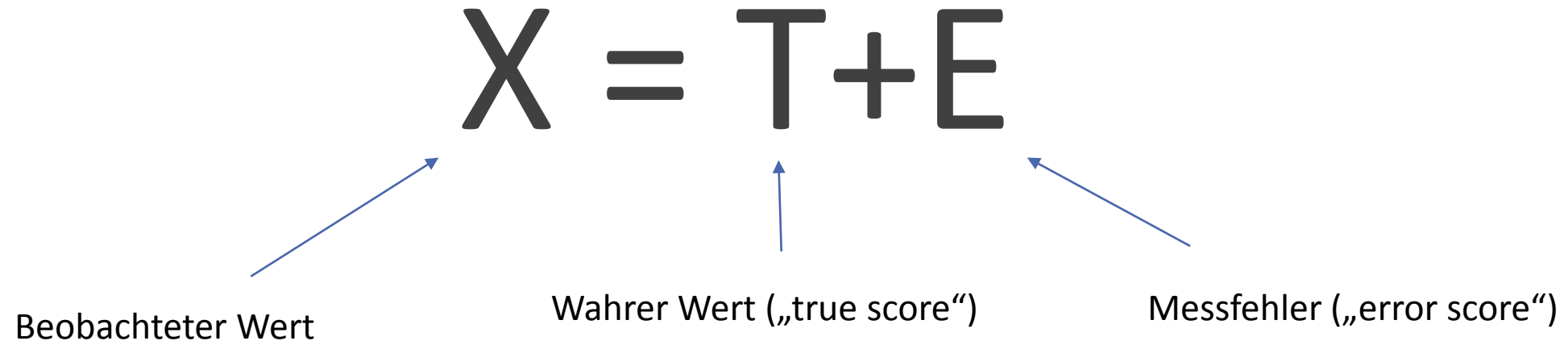
Beispiel 2: IQ-Messung

- Wert 1. Termin: IQ = 135
- Wert 2. Termin: IQ = 119
 - hat sich die Intelligenz der Person verändert?

Klassische Testtheorie - Annahmen

$$X = T + E$$

Beobachteter Wert Wahrer Wert („true score“) Messfehler („error score“)

The diagram shows the equation $X = T + E$ in large, bold, black font. Below the equation, three labels are positioned: 'Beobachteter Wert' on the left, 'Wahrer Wert („true score“)' in the center, and 'Messfehler („error score“)' on the right. Three blue arrows point from each label to its corresponding variable in the equation: one from 'Beobachteter Wert' to 'X', one from 'Wahrer Wert („true score“)' to 'T', and one from 'Messfehler („error score“)' to 'E'.

Klassische Testtheorie - Annahmen

- der wahre Wert einer Person ist bei jeder Messung gleich
- der Messfehler variiert von Messung zu Messung
- der Messfehler ist unabhängig vom wahren Wert
- die Messfehler zweier Tests A und B hängen nicht zusammen
- die Messfehler in einem Test A sind unabhängig von den wahren Werten in Test B

Klassische Testtheorie - Annahmen

Beispiel Hochsprung:

- der wahre Wert einer Person ist bei jeder Messung gleich
→ die beobachteten Werte (Sprunghöhen) schwanken

Klassische Testtheorie - Annahmen

Beispiel Hochsprung:

- der Messfehler variiert von Messung zu Messung
→ z.B. unterschiedliche Einflüsse von Motivation, Müdigkeit, Wind...

Klassische Testtheorie - Annahmen

Beispiel Hochsprung:

- der Messfehler ist unabhängig vom wahren Wert

Klassische Testtheorie - Annahmen

Beispiel Hochsprung:

- die Messfehler zweier Tests A und B hängen nicht zusammen

Klassische Testtheorie - Annahmen

Beispiel Hochsprung:

- die Messfehler in einem Test A sind unabhängig von den wahren Werten in Test B
 - Schwankungen in anderen Merkmalen haben keine Auswirkungen auf die Genauigkeit der Messungen der Sprunghöhe

Reliabilität

- Messgenauigkeit eines Tests
- Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte

Arten der Reliabilitätsbestimmung

- Retest-Reliabilität
- Paralleltest- Reliabilität
- Testhalbierungs- Reliabilität
- Interne Konsistenz

Retest-Reliabilität

- Testwiederholung
- gleicher Test zu zwei Zeitpunkten an derselben Stichprobe durchgeführt
- Bestimmung der Korrelation zwischen beiden Messzeitpunkten

Paralleltest-Reliabilität

- zwei inhaltlich äquivalente Tests zu zwei Zeitpunkten an derselben Stichprobe gemessen
- Korrelation der beiden Tests

Testhalbierungs-Reliabilität

- einmalige Testdurchführung an einer Stichprobe
- Aufteilung des Tests im Nachhinein in zwei möglichst gleiche Hälften
- Bestimmung der Korrelation zwischen den beiden Testhälften

Interne Konsistenz

- Test wird in so viele Teile zerlegt, wie es Items gibt
- nur sinnvoll möglich bei homogenen Tests
- Schätzung der durchschnittlichen Korrelation zwischen den Items unter Berücksichtigung der Testlänge

Einflussfaktoren auf die Reliabilität

- Homogenität /Heterogenität der Items
- Varianz
- Messfehler

Wie hoch sollte die Reliabilität sein?

- abhängig vom Merkmal
- Vergleich mit ähnlichen Verfahren
- Abhängig von Einsatzbedingungen
- Kosten-Nutzen-Abwägung
- möglichst hoch

Wie kann man die Reliabilität verbessern?

- Homogenität erhöhen (wenn sinnvoll)
- Testverlängerung
- Messfehler reduzieren
 - durch klare Instruktion
 - durch klare Formulierung der Items
 - durch standardisierte Testbedingungen

Kritik an der klassischen Testtheorie

Vorzüge:

- Sparsam
- Wenige Grundannahmen
- Ableitung der Reliabilität

Kritikpunkte:

- Axiome empirisch nicht prüfbar
- Keine zufällige Verteilung der Messfehler um den wahren Wert
- Zweifel an Unkorreliertheit von wahren Wert und Messfehler

Ausblick

Nächste Sitzung: Gütekriterien II

Vielen Dank für die
Aufmerksamkeit!

Literatur

- Bortz, J., & Döring, N. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5. Auflage. Heidelberg: Springer-Verlag → Kap. 3.2, Kap. 3.3
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. 3. Auflage. München: Pearson Studium → Kap. 2, Kap. 4.3
- Hussy, W., Schreier, M., & Echterhoff, G. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften-für Bachelor*. Berlin, Heidelberg: Springer-Verlag. → Kap. 8.1
- Moosbrugger, H., & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*. 2. Auflage. Berlin, Heidelberg: Springer-Verlag. → Kap. 2
- Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische Diagnostik*. 5. Auflage. Berlin, Heidelberg: Springer-Verlag. → Kap. 2.1.2