

Testtheorie und Gütekriterien von Messinstrumenten

Objektivität

Reliabilität

Validität

Genauigkeit von Messungen

- ◆ Jede Messung zielt darauf ab, möglichst exakte und fehlerfreie Messwerte zu erheben.
- ◆ Dennoch ist jeder Messvorgang (mehr oder weniger stark) mit Messfehlern behaftet.
- ◆ Um Aussagen über die Genauigkeit von Messungen zu ermöglichen, wird eine statistische Theorie zur Entstehung von Messfehlern benötigt.
- ◆ Hierfür bildet die sogenannte **„klassische Testtheorie“** die Grundlage.

Klassische Testtheorie

Grundgleichung

- ◆ Die klassische Testtheorie geht davon aus, dass sich ein beobachteter (gemessener) Wert (observed score) aus der Summe eines ‚wahren Wertes‘ (true score) und einem Messfehler (random error) zusammensetzt.
- ◆ Grundgleichung:

Beobachteter Wert (X) = Wahrer Wert (T) + Messfehler (e)

$$X = T + e$$

Klassische Testtheorie

Axiome

1. Der Mittelwert der Messfehler ist gleich Null.
2. Die Messfehler korrelieren nicht mit den wahren Werten der Messung.
3. Die Messfehler zweier Messwertreihen sind unkorreliert.
4. Die Messfehler einer Messwertreihe korrelieren nicht mit den wahren Werten einer anderen Messung.

Objektivität

- ◆ Eine Messung entspricht dem Gütekriterium der **Objektivität**, wenn zwei Anwender mit dem gleichen Messinstrument übereinstimmende Resultate erzielen. Als Maß für die Objektivität kann der Korrelationskoeffizient für das Ausmaß der Übereinstimmung herangezogen werden.
- ◆ Man kann drei Formen der Objektivität unterscheiden:
 - **Durchführungsobjektivität**
 - **Auswertungsobjektivität**
 - **Interpretationsobjektivität**

Formen der Objektivität

- ◆ Die **Durchführungsobjektivität** ist beeinträchtigt, wenn zwei Anwender bei Verwendung desselben Messinstruments ein unterschiedliches Antwortverhalten beim Befragten evozieren.
- ◆ Die **Auswertungsobjektivität** ist dann beeinträchtigt, wenn zwei Auswerter bei gleichem Antwortverhalten zu unterschiedlichen Messergebnissen kommen.
- ◆ Die **Interpretationsobjektivität** ist dann beeinträchtigt, wenn zwei Interpretatoren aus den gleichen Auswertungsergebnissen unterschiedliche Schlüsse ziehen.

Reliabilität

- ◆ Die Reliabilität bzw. Zuverlässigkeit einer Messung bezeichnet das Ausmaß mit dem wiederholte Messungen mit dem gleichen Messinstrument die gleichen Messwerte liefern, wobei das Ziel darin besteht, dass der Messwert (beobachtete) Wert möglichst genau mit dem wahren Wert der Messung übereinstimmt.
- ◆ Da der wahre Wert einer Messung nicht bekannt ist, kann die Reliabilität nur geschätzt werden.

Verfahren zur Schätzung der Reliabilität

- ◆ Zur Schätzung der Reliabilität können drei Verfahren angewendet werden:
 - **Test-Retest-Methode** – das Messinstrument wird nach einem Zeitintervall erneut angewendet. Die Korrelation der Messwerte zum Zeitpunkt A mit den Messwerten zum Zeitpunkt B informiert über die Test-Retest-Reliabilität.
 - **Paralleltestmethode** – die Messung erfolgt mit zwei vergleichbaren Messinstrumenten. Die Korrelation der Messwerte des Instruments A mit den Messwerten des Instruments B stellt eine Schätzung der Paralleltest-Reliabilität dar.
 - **Splithalf-Methode** – Messinstrumente mit multiplen Indikatoren werden in zwei Hälften aufgeteilt. Die Korrelation zwischen den beiden Testhälften gibt Aufschluss über die Splithalf-Reliabilität.

Validität

- ◆ Die Validität bzw. Gültigkeit eines Messinstruments bezeichnet das Ausmaß, in dem das Messinstrument tatsächlich das misst (z.B. Persönlichkeitsmerkmal, Verhaltensweise, Einstellung), was es zu messen vorgibt.
- ◆ Man kann drei verschiedene Formen der Validität unterscheiden:
 - Inhaltsvalidität
 - Kriteriumsvalidität
 - Konstruktvalidität

Formen der Validität

- ◆ **Inhaltsvalidität** – ein inhaltsvalides Messinstrument liegt dann vor, wenn die ausgewählten Items die zu messende Eigenschaft in hohem Maße repräsentieren.
- ◆ **Kriteriumsvalidität** – wird durch die Korrelation der Messwerte mit einem Außenkriterium berechnet, wobei das Außenkriterium unabhängig von der eigentlichen Messung mit einem anderen Messinstrument erhoben werden muss. Man unterscheidet hier zwischen **Übereinstimmungsvalidität** und **Vorhersagevalidität**.
- ◆ **Konstruktvalidität** – liegt dann vor, wenn das von einem Messinstrument erfasste Konstrukt mit möglichst vielen anderen Variablen in theoretisch begründbaren Zusammenhängen steht. Hieraus müssen Hypothesen ableitbar sein, die einer empirischen Überprüfung standhalten.

Verhältnis der Gütekriterien

- ◆ Die drei Gütekriterien – Objektivität, Reliabilität und Validität – stehen in einem hierarchischem Verhältnis zueinander:
 - Objektivität ist eine notwendige, aber nicht hinreichende Bedingung der Reliabilität.
 - Reliabilität ist eine notwendige, aber nicht hinreichende Bedingung der Validität.
 - Das Ziel ist die Konstruktion valider Messinstrumente.