

# Gütekriterien

**Erziehungs- und Sozialwissenschaften  
«Beurteilung & Förderung»**

Irene Althaus



# Was ist ein guter Test?

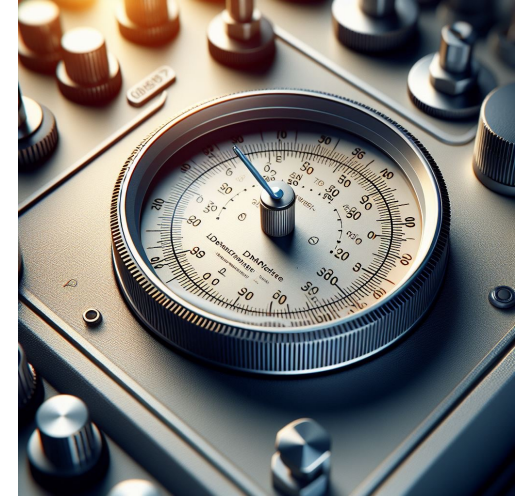
Ziel: Leistungsbewertung so zu erstellen, dass mit ihnen gut gemessen werden kann.

1. Tests müssen Gütekriterien folgen. Was genau bedeutet das?
2. Wann ist ein Test gut?
3. Wie kann bereits bei der Testerstellung dafür Sorge getragen werden, dass ein Test die Gütekriterien erfüllt?
4. Wie kann die Lehrperson sicher stellen, dass Sie nach den Gütekriterien beurteilt?

# Klassische Testtheorie

Eine gute Ausgangsbasis zur Herleitung der Eigenschaften eines guten Tests bietet die klassische Testtheorie (Bortz, 1995).

- ➔ **Maximierung der Genauigkeit** einer Messung
- ➔ **Minimierung des jeweiligen Messfehlers.**



# Gütekriterien

## Kriterien für die Güte eines Tests

(Arnold 2002; Bortz/Döring 1995; Heller/Hany 2002; Weinert 2002).

Ein guter Test ...

- ermittelt verteilte Ergebnisse, lässt Qualitätsaussagen zu.
- hat einen angemessenen Schwierigkeitsgrad.
- trennt die guten von den schlechten Leistungen.
- Ist, wenn die einzelnen Aufgaben trennscharf sind.
- ist **objektiv**.
- ist **reliabel**.
- ist **valide**.



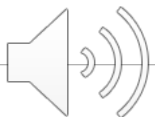
# Objektivität

Eine Klassen- / Schularbeit sind objektiv, wenn ihre Durchführung und ihre Auswertung von der jeweiligen Lehrperson, die dafür verantwortlich ist, unabhängig sind (Grotjahn, 2006, S.222), d.h. wenn **subjective judgement** ausgeschlossen werden kann (Hughes, 2003, S.22)



Bild erstellt mit bingcreate.com

- ➔ Objektivität bedeutet bei der Durchführung, Auswertung und Interpretation von Tests, Subjektivität zu reduzieren.



# Objektivität

Dlaska & Krekeler (2009):  
“Gerechtigkeit” bei informeller Leistungsbeurteilung

4 Bereiche

1. Ausreichende Transparenz
2. Schlüssige Bewertung
3. Konsistente und übertragbare Leistungsmessungen
4. Gültige Interpretation

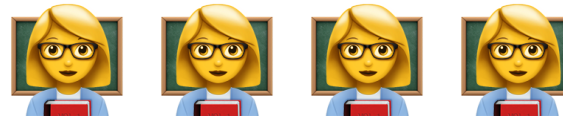
→ Umfassen die weiteren Testgütekriterien: Reliabilität und Validität



# Reliabilität



## Intrarater-Reliabilität



## Interrater-Reliabilität



# Reliabilität

1. Geschlossene und offene Aufgabenformate unterstützen Reliabilität in unterschiedlichem Ausmass
2. Innere Konsistenz von Testitems ist Voraussetzung für das Messen eines Konstrukts
3. Testlänge, homogene Items und trennscharfe Aufgaben erhöhen die Reliabilität

→ Reliabilität im schulischen Kontext: Klare und Transparente Angaben hinsichtlich Lern- und Bewertungszielen



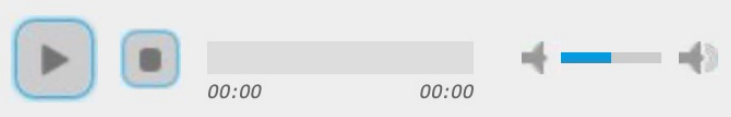


# Validität

„Die Validität bezieht sich darauf, inwieweit ein Test das erfasst, was er erfassen soll, und inwieweit er zu fairen Entscheidungen führt.“  
(Grotjahn, 2006, S. 223)

Die Validität gibt als wichtigstes Gütekriterium eines schulischen Tests an, ob ein Test auch das misst, was er messen soll (Bortz 1995; Moosburger/ Kelava 2007).

? Test your Level of English Listening



Question 1 of 15

What kind of film is the woman talking about?

- ☐ a horror film
- ☐ a thriller
- ☐ a historical drama
- ☐ a romantic comedy



# Literatur

# Literatur

- Arnold, K.-H. (2002): Qualitätskriterien für die standardisierte Messung von Schulleistungen. In: F. E. Weinert (Hrsg.), Leistungsmessungen in Schulen. Weinheim: Beltz, 117–130.
- Bortz, J. (1995): Lehrbuch der empirischen Forschung. Berlin: Springer.
- Bortz, J. / Döring, N. (1995): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin: Springer.
- Diakka, A. & Krekeler, Ch. (2009). Sprachtestes. Leistungsbeurteilungen im Fremdsprachenunterricht evaluieren und verbessern. Baltmannsweiler: Schneider.
- Grotjahn, R. (2006). Prüfen – Testen – Bewerten. In: Jung, U.O.H. (Hrsg.), 221-230
- Heller, K. A. / Hany, E. A. (2002): Standardisierte Schulleistungsmessungen. In: F. E. Weinert (Hrsg.), Leistungsmessungen in Schulen. Weinheim: Beltz, 87–102.
- Hinger, B. & Stadler, W. (2018). Testen und Bewerten fremdsprachlicher Kompetenzen. Tübingen: Narr Francke Attempto Verlag.
- Hughes, A. (2003). Testing for Language Teachers. Cambridge: Cambridge University Press.
- Lerche, T. (2022). Leistungsbeurteilung an Schulen. München: Ludwig-Maximilians-Universität.
- Weinert, F. E. (2002): Vergleichende Leistungsmessungen in Schulen – eine umstrittene Selbstverständlichkeit. In: F. E. Weinert (Hrsg.), Leistungsmessungen in Schulen. Weinheim: Beltz, 17–32.