

Methoden der Sportwissenschaft 1

Arbeitsblatt 1.3: Gütekriterien empirischer Erhebungsverfahren

Gütekriterien empirischer Erhebungsverfahren

Hauptgütekriterien

Die **Objektivität** ist der Grad, in dem die Testergebnisse unabhängig vom Untersucher sind. Man unterscheidet: die Durchführungsobjektivität, die Auswertungsobjektivität und die Interpretationsobjektivität

Die **Reliabilität** ist das Ausmaß der Messgenauigkeit eines Tests. Wichtig ist die Unterscheidung von Test-Retest-Reliabilität (Stabilität des Merkmals) und Paralleltestreliabilität (Äquivalenz des Testverfahrens). Für die Beurteilung der Reliabilitätskoeffizienten gelten folgende Richtwerte: >0.90 = ausgezeichnet; $0.80 - 0.90$ = sehr gut; $0.70 - 0.80$ = annehmbar; $0.60 - 0.70$ = mäßig; < 0.60 = gering (Bös, 1987).

	Reliabilitätsarten		
	Stabilität	Äquivalenz	Konsistenz
Test-Retest-Methode	X		
Paralleltest-Methode		X	
Split-Half-Methode			X
Konsistenz-Analyse			X

Abbildung 1: Methoden zur Bestimmung der Reliabilität

Die **Validität** eines Tests gibt die Genauigkeit an, mit der der Test dasjenige Merkmal, das er erfassen soll, tatsächlich auch misst. Man unterscheidet im Wesentlichen folgende Validitätsarten: die inhaltlich-logische Validität, die Expertenvalidität, die Kriteriumsvalidität (Vergleich mit einem Außenkriterium) und die Konstrukt oder faktorielle Validität (Faktorenladung eines Tests)

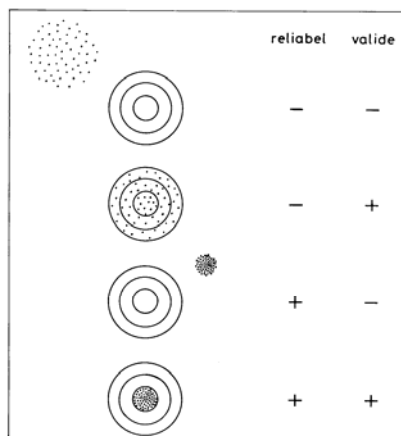


Abbildung 2: Zusammenhang zwischen Reliabilität und Validität

Nebengütekriterien

Ein Test ist **ökonomisch**, wenn er hinsichtlich der organisatorischen, räumlichen, zeitlich/personellen, instruktions- und gerätespezifischen Testdurchführungsbedingungen keine oder nur geringe Ansprüche an Testleiter und Testpersonen stellt.

Ein Test ist **nützlich**, wenn für den Testinhalt ein praktisches Bedürfnis besteht, und es die Testanwendung erlaubt, relevante Entscheidungen zu treffen.

Ein Test ist **normierbar**, wenn die Messwerte einer Versuchsperson in Bezug gesetzt werden können zu den Testergebnissen einer ausgewählten Population. Die Rohwerte werden dazu in sogenannte Normwerte umgewandelt (siehe Abbildung 3)

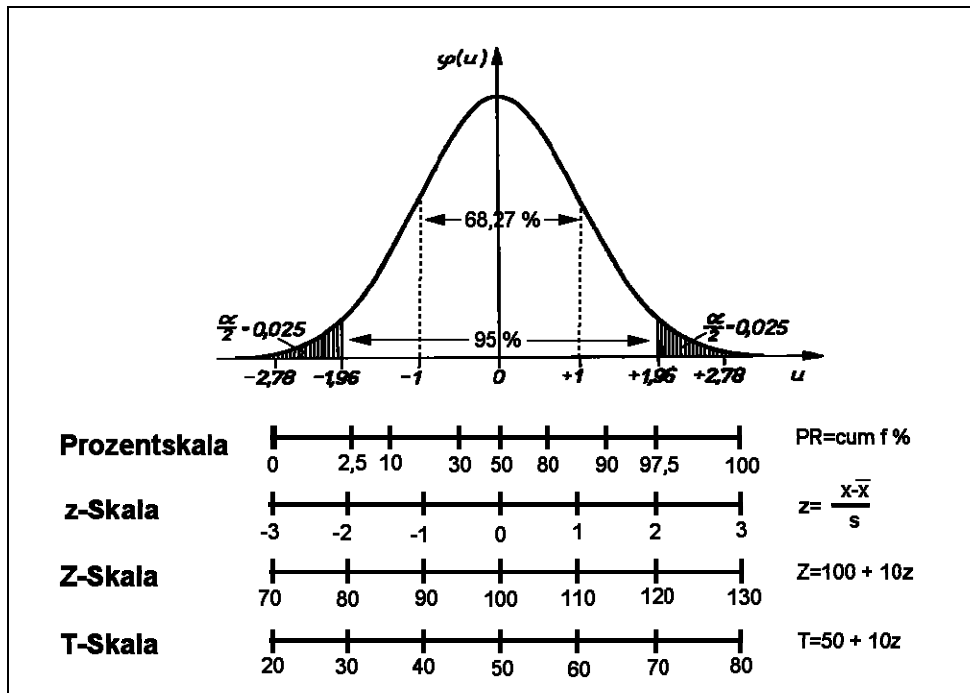


Abbildung 3: Wichtige Normwerttransformationen

Israel (1983) unterscheidet des weiteren verschiedene Normbegriffe: **Majoritätsnormen** (Orientierung an der Standardnormalverteilung); **Minimalnormen** (Abgrenzung zu pathologischen Werten); **Idealnormen** (im Hinblick auf die Gesundheit und die Lebenserwartung ideal); **Spezialnormen** (Voraussetzungen für bestimmte Sportarten, z. B. Körperfettanteil) (Wydra, 2004).

Die **Schwierigkeit** eines Tests gibt an, inwieweit er interindividuelle Unterschiede feststellen kann. Sie ist abhängig von der Aufgabenschwierigkeit und der Differenziertheit der Skalierung. Im Allgemeinen besteht zwischen Trennschärfe und Schwierigkeit ein parabolischer Zusammenhang, d. h. bei einer mittleren Schwierigkeit erzielt man die größte Trennschärfe.

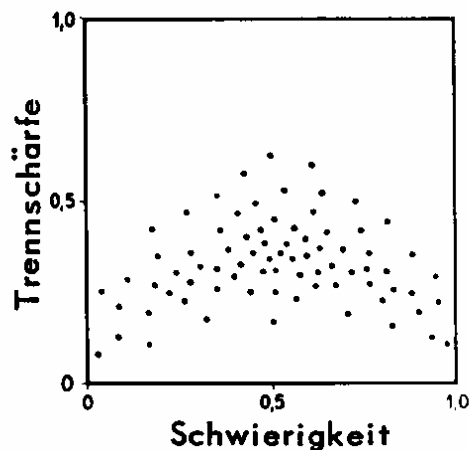


Abbildung 4: Zusammenhang zwischen Trennschärfe und Schwierigkeit eines Testverfahrens

Die Bestimmung der Hauptgütekriterien

Die Bestimmung von Objektivität, Test-Retest-Reliabilität und Kriteriums-Validität erfolgt im Allgemeinen über den Korrelationskoeffizienten (vgl. Bös, 1987). Der Korrelationskoeffizient ist ein Maß für den Zusammenhang zwischen zwei Variablen. Er kann Werte zwischen +1 und -1 annehmen.

Für die Beurteilung der Reliabilitätskoeffizienten gelten folgende Richtwerte: >0.90 = ausgezeichnet; $0.80 - 0.90$ = sehr gut; $0.70 - 0.80$ = annehmbar; $0.60 - 0.70$ = mäßig; < 0.60 = gering (Bös, 1987).

Korrelationen können graphisch bestimmt werden (siehe Abbildung 1 und Abbildung 2). Die genaue Bestimmung der Koeffizienten erfolgt über die Berechnung des Korrelationskoeffizienten. Eine Testevaluation eines Bauchmuskeltests ist bei Wydra (1995) dargestellt.

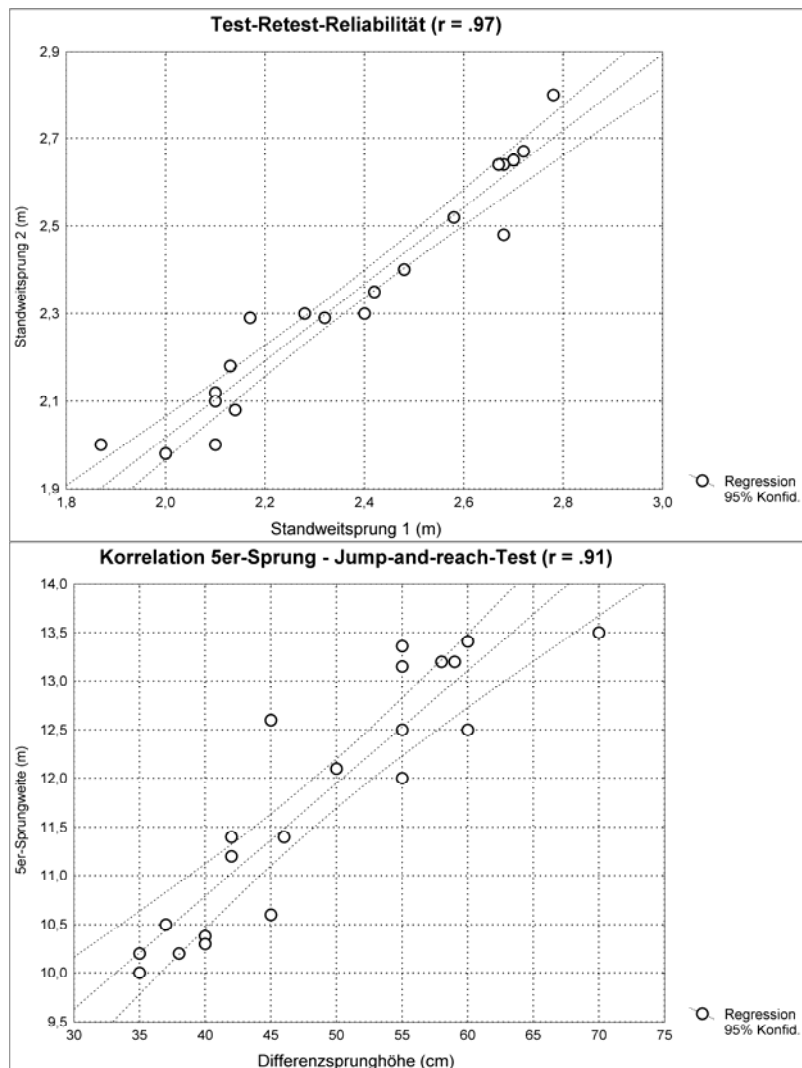


Abbildung 5: Beispielhafte zeichnerische Darstellung des Test-Retest-Koeffizienten bei der Testübung „Standweitsprung“ zur Erfassung der Sprungkraft.

Abbildung 6: Beispielhafte zeichnerische Darstellung der Kriteriumsvalidität der Testübung „Standweitsprung“ über die Korrelation mit dem Jump-and-reach-Test.

Besondere Hinweise

Was passiert mit dem Korrelationskoeffizienten zwischen zwei sportmotorischen Tests, wenn die Untersuchung desselben

- a. bei einer homogenen Gruppe männlicher Sportstudenten,
- b. bei einer heterogenen Gruppe gemischtgeschlechtlicher Freizeitsportler durchgeführt wird? Warum ist dies so?

Mit einem Beispiel soll versucht werden, das Problem zu erläutern. Bestimmt wird der korrelative Zusammenhang zwischen zwei Ausdauer tests. In diesem Fall wurde die Korrelation zwischen der Laufleistung im 6-Minuten-Lauf und der Leistung bei der Fahrradergometrie bestimmt.

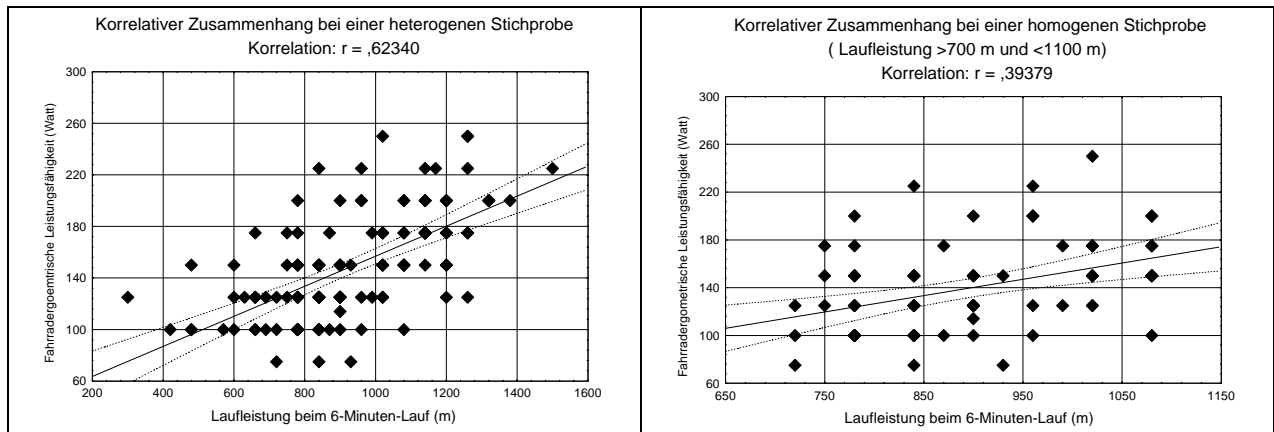


Abbildung 7: Graphische Darstellung des Einflusses der Stichprobenstreuung auf die Höhe des Korrelationskoeffizienten. Links die Korrelation in der Gesamtstichprobe (n=128), rechts die Korrelation in der Teilstichprobe (n=78), wobei nur die Vpn. berücksichtigt wurden, deren Messwerte im Bereich $M \pm SD$ lagen.

Der Korrelationskoeffizient ist in der heterogenen Stichprobe höher als in der homogenen Stichprobe. Das hängt mit der Streuung der Messwerte zusammen.

FORMEL (ii): PRODUKT-MOMENT-KORRELATIONSKOEFFIZIENT

$$r_{xy} = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{(N \sum x_i^2 - (\sum x_i)^2) (N \sum y_i^2 - (\sum y_i)^2)}}$$

Abbildung 8: Formel zur Berechnung des Produkt-Moment-Korrelationskoeffizienten (Bös, 1986, S. 49)

Im Zähler der Formel wird die Summe der Quadrate der Abweichungen der einzelnen Messwerte vom jeweiligen Mittelwert berechnet. Bei Messreihen aus einer heterogenen Stichprobe resultieren daraus größere Abweichungen und damit auch größere Produktsummen.

Zum Teil kann eine Erhöhung des Korrelationskoeffizienten mit dem Problem der sog. „Scheinkorrelation“ zusammenhängen. Hierbei kommt es durch das Zusammenfügen von zwei oder mehr leistungsdifferenten Gruppen, in denen relativ geringe oder keine Korrelationen zu beobachten sind, zu einer Gesamtgruppe zu einer relativen hohen Korrelation.

Literatur:

- Israel, S. (1983). Körperliche Normbereiche in ihrem Bezug zur Gesundheitsstabilität. *Theorie und Praxis der Körperkultur*, 32,, 360 - 363.
- Bös, K. (1986). Statistikkurs I (3. Aufl.) Ahrensburg: Czwalina.
- Bös, K. (1987). *Handbuch sportmotorischer Tests*. Göttingen: Hogrefe.
- Bös, K. (Hrsg.) (2001). *Handbuch motorische Tests*. Göttingen: Hogrefe.
- Wydra, G. (1995). Ein neuer Test zur Beurteilung der Kraft der Bauchmuskulatur. *Krankengymnastik*, 47, 937 - 946. (Verfügbar unter: <http://www.uni-saarland.de/fak5/sportpaed/tests/crunches.pdf>)
- Wydra, G. (2004). Zur Problematik von Normwerten in der Bewegungstherapie. *Krankengymnastik – Zeitschrift für Physiotherapeuten*, 56, 2280 – 2289.
- Wydra, G. (2006). Normierung der motorischen Leistungsfähigkeit. *Bewegungstherapie und Gesundheitssport*, 22, 223 - 227.