



Einführung in die Forschungsmethoden der Psychologie

BSc Philosophie-Neurowissenschaften-Kognition WiSe 23/24

BSc Psychologie WiSe 23/24

Prof. Dr. Dirk Ostwald

(11) Validität

Einordnung und frühe Perspektiven

Konstruktvalidität

Argumentatorische Validität

Kausale Validität

Selbstkontrollfragen

Einordnung und frühe Perspektiven

Konstruktvalidität

Argumentatorische Validität

Kausale Validität

Selbstkontrollfragen

Klassische Gütekriterien für Leistungs- und Persönlichkeitstests

Objektivität

- Unabhängigkeit des Testergebnisses von Durchführenden und Auswertenden

Reliabilität

- Replizierbarkeit des Testergebnisses

Validität

- Grad der Gültigkeit des Testergebnisses bezüglich des Testgegenstandes

Einordnung

- Der Validitätsbegriff und seine Literatur entstammen den Humanwissenschaften
- Validität dient als Autoritätsgrundlage von psychozialen Messungen
- Beispiele psychosozialer Messungen: Leistungs-, Persönlichkeits- und Einstellungsmessungen
- Der Validitätsbegriff ist der für die psychologische Anwendungsforschung relevante Begriff
- Die Literatur zum Validitätsbegriff ist von der messtheoretischen Literatur getrennt
- Der Messbegriff wird bei psychosozialer Messungen nicht formal-philosophisch betrachtet
- Insbesondere ignoriert die Psychologie die für sie geschaffene Repräsentationstheorie
- Der Validitätsbegriff ist historischen Wandlungen unterworfen
- Es gibt bis heute keinen Konsens zum Validitätsbegriff des Messens

Frühe Perspektiven und Operationalisierung

Validität als “der Grad mit dem ein Test das misst, was er zu messen vorgibt”

Kriterienvalidität (vgl. Kelley (1927), Cureton (1950))

- “Prädiktive Validität”
- Validität als Korrelation von Testergebnissen zwischen Tests (“Validitätskoeffizienten”)
- Bspw. Korrelation von Studieneignungstestscores mit Psychologiestudiumnoten

Inhaltsvalidität (vgl. Lindquist (1951), Sireci (1998))

- Validität als adäquate Abdeckung der Testdomäne
- Auswahl von Testaufgaben anhand von Expert:inneneinschätzungen und Dokumentation
- Bspw. Inhaltsvalidität von Klausuren im Studium

Einordnung und frühe Perspektiven

Konstruktvalidität

Argumentatorische Validität

Kausale Validität

Selbstkontrollfragen

Cronbach und Meehl (1955) Construct Validity in Psychological Tests

Motivation

- APA Committee on Psychological Tests (1950 - 1954)
- Welche Gütekriterien sollte ein psychologischer Test vor seiner Veröffentlichung erfüllen?
- Zeitgenössische unbefriedigende Vielfachheit von Validitätskriterien
- Kriterienvalidität und Inhaltsvalidität oft nicht adäquat, z.B. Persönlichkeitstests

Definition des Konstruktbegriffs

- "A postulated attribute of people, assumed to be reflected in test performance (...). A construct is defined implicitly by a network of associations and propositions in which it occurs (...) Constructs may vary in nature from those very close to "pure description" (...) to highly theoretical constructs involving hypothesized entities and processes (...)."

Definition der Konstruktvalidität

- "Many types of evidence are relevant to construct validity, including content validity, interitem correlations, intertest correlations, test-"criterion" correlations, studies of stability over time and stability under experimental intervention (...). Construct validity cannot generally be expressed in the form of a single simple coefficient."

Cronbach und Meehl (1955) Construct Validity in Psychological Tests

Konstruktvalidierung

- Artikulation der konstruktimmanenten theoretischen Konzepte und ihrer Relationen
- Entwicklung von Messmethoden, die die theoretischen Konzepte abbilden
- Empirische Überprüfung der postulierten Relationen der theoretischen Konzepte

Beispiel einer Konstruktvalidierungsprozedur

Einem Studenten werde mitgeteilt, dass er durch eine wichtige Prüfung gefallen ist.

Es korreliere dann Persönlichkeitstest X mit $\rho = .5$ mit der Stärke des Handschweißes.

Der Test X erfüllt dann eine Kriterienvalidität (Prädiktive Validität).

Wenn dann postuliert wird Test X messe "Angstneigung", sind auch alternative Interpretationen möglich.

Beispielsweise könnte Test X auch "Akademischen Ehrgeiz" messen.

Weitere Studien könnten zeigen, dass Test X mit Angespanntheitsratings, intellektueller Einschränkung bei Elektroschocks und der Taylor Anxiety Scale korreliert, dass X besonders hohe Werte bei Angststörungspatient:innen, und mit Angst vor dem Durchfallen in einer anderen Prüfung korreliert. Andere Studien könnten zeigen, dass X nicht mit sozialer Klasse, Berufsziel oder Wertorientierung korrelieren, so dass "Akademischer Ehrgeiz" eher als das durch X gemessene Konstrukt nicht in Frage kommt.

Campbell and Fiske (1959) Convergent and discriminant validation

Konvergente Validität

- Konvergenz von Messungen der gleichen Eigenschaft durch verschiedene Messinstrumente
- Wenn Test X Extraversion misst, sollte Test X mit Fremdeinschätzung korrelieren

Divergierende Validität

- Divergenz von Messungen von verschiedenen Eigenschaften durch ähnliche Messinstrumente
- Wenn Text X Extraversion und Test Y Dominanz misst, sollten X und Y nicht stark korrelieren.

Einordnung und frühe Perspektiven

Konstruktvalidität

Argumentatorische Validität

Kausale Validität

Selbstkontrollfragen

Messick (1993) Foundations of validity

- Verschiedene Formen von Validität als verschiedene Formen der Evidenz für einen Test.
- Validität eines Tests als evidenzbasierte epistemische Aussage einer Wissenschaftscommunity.
- Validierung von Tests sollte die Konsequenzen und den Gebrauch von Tests berücksichtigen.
- Einflussreiche Validitätskonzeption im Bereich der Bildungsforschung.

Beispiel

- Wenn Schultests als Grundlage für die Beschäftigung von Lehrer:innen dienen, sollte nicht nur gezeigt werden, dass sie das Wissen und die Fähigkeiten von Schüler:innen messen, sondern auch, dass die auf ihnen basierte Entscheidungen hinsichtlich der Lehrkräfte die gewünschten positiven Effekte haben und keine unerwarteten negativen Effekte.

Kane (1992) An argument-based approach to validity

- Testvalidierung als Argumentkonstruktion zur Verteidigung eines Tests für einen Zweck.
- Pragmatische, kontext- und publikumsabhängige Konzeption der Validität.
- Fokus auf der Funktion eines Tests für einen Zweck, nicht der Messung einer Eigenschaft.

AERA, APA, NCME (2014) Standards for Educational and Psychological Testing

“Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests. The process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations. It is the interpretations of test scores for proposed uses that are evaluated, not the test itself. When test scores are interpreted in more than one way (e.g., both to describe a test taker’s current level of the attribute being measured and to make a prediction about a future outcome), each intended interpretation must be validated. Statements about validity should refer to particular interpretations for specified uses. It is incorrect to use the unqualified phrase “the validity of the test.””

⇒ Tests als Messinstrumente nur eine der möglichen Funktionen von Tests.

Einordnung und frühe Perspektiven

Konstruktvalidität

Argumentatorische Validität

Kausale Validität

Selbstkontrollfragen

Borsboom (2004) The concept of validity

Ein Test ist ein valides Messinstrument einer Eigenschaft, wenn

- (a) die Eigenschaft existiert
- (b) Variabilität der Eigenschaft Variabilität der Testergebnisse induziert

Kein Forschungsfokus auf die Relationen zwischen Eigenschaften

Forschungsfokus auf Prozesse, die eine Eigenschaft in Testresultaten abbilden

⇒ Historische Wandlung des Validitätsbegriffs

- Realismus (“Ein valider Test misst, was er zu messen vorgibt”)
- Logischer Positivismus (Operationalistische Konstruktvalidität)
- Pragmatismus (Adäquatheit für einen Zweck)
- Prozessmodell-basierter Realismus

Selbstkontrollfragen

1. Nennen und erläutern Sie die klassischen Gütekriterien für Leistungs- und Persönlichkeitstests.
2. Erläutern Sie den Begriff der Kriterienvalidität nach Kelley (1927).
3. Erläutern Sie den Begriff der Inhaltsvalidität nach Lindquist (1951).
4. Erläutern Sie den Begriff des Konstrukts nach Cronbach and Meehl (1955).
5. Erläutern Sie den Begriff der Konstruktvalidität nach Cronbach and Meehl (1955).
6. Nennen Sie drei Schritte zur Konstruktvalidierung nach Cronbach and Meehl (1955).
7. Erläutern Sie die Begriffe der konvergenten und divergierenden Validität nach Campbell and Fiske (1959).
8. Erläutern Sie den Begriff der Argumentatorischen Validität nach Messick (1993) und Kane (1992).
9. Erläutern Sie den Begriff der Kausalen Validität nach Borsboom, Mellenbergh, and Van Heerden (2004).

- AERA, American Educational Research Association, American Psychological Association APA, and National Council on Measurement in Education NCME. 2014. *Standards for Educational and Psychological Testing*. Washington, D.C: American Educational Research Association.
- Borsboom, Denny, Gideon J. Mellenbergh, and Jaap Van Heerden. 2004. "The Concept of Validity." *Psychological Review* 111 (4): 1061–71. <https://doi.org/10.1037/0033-295X.111.4.1061>.
- Campbell, Donald T, and Donald W Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56 (2).
- Cronbach, Lee J, and Paul E. Meehl. 1955. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52 (4).
- Cureton, Edward E. 1950. "Validity, Reliability, and Baloney." *Educational and Psychological Measurement* 10 (1): 94–96. <https://doi.org/10.1177/001316445001000107>.
- Kane, Michael T. 1992. "An Argument-Based Approach to Validity." *Psychological Bulletin* 112 (3): 527–35.
- Kelley, Truman. 1927. *Interpretation of Educational Measurements*. Yonkers-on-Hudson.
- Lindquist, M. 1951. *Educational Measurement*. Washington, DC: American Council on Education.
- Messick, Samuel. 1993. "Foundations of Validity: Meaning and Consequences in Psychological Assessment." *ETS Research Report Series* 1993 (2). <https://doi.org/10.1002/j.2333-8504.1993.tb01562.x>.
- Sireci, Stephen G. 1998. "The Construct of Content Validity." *Social Indicators Research* 45 (1/3,): 83–117. <https://www.jstor.org/stable/27522338>.