

DGFF-Kolleg

Auswahlbibliographie zum Thema
Tests in der Fremdsprachenforschung



23. März 2022

Claudia Harsch

harsch@uni-bremen.de



Empfohlene Literatur zum Einstieg

Bachman, Lyle F. (2004): *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.

Bachman, Lyle F. & Kunnan, Anthony J. (2005): *Statistical analyses for language assessment. Workbook and CD*. Cambridge: Cambridge University Press.

Diese beiden Bände geben eine sehr gute Einführung in die statistische Testanalyse. Das Workbook (mit CD) ergänzt die Monographie um praktische Beispiele und Übungen an realen Datensätzen.

Bachman, Lyle F. & Palmer, Adrian S. (2010): *Language assessment in practice*. Oxford: Oxford University Press.

Diese Monographie ist eines der Standardwerke in der Testliteratur; sie diskutiert alle wesentlichen Aspekte des Designs, der Entwicklung und des Nutzens von Sprachtestes und Sprachbeurteilung. Insbesondere die Ausführungen zum Assessment Use Argument sind bemerkenswert, da sie die Testnutzung und den Einsatz von Beurteilung in den Mittelpunkt rücken.

Douglas, Dan (2010): *Understanding Language Testing*. London: Hodder Education.

Dieser Band bietet eine kurze und leicht verständliche Einführung in die Natur, Entwicklung, Analyse und den Einsatz von Sprachtests.

Hinger, Barbara & Stadler, Wolfgang (2018): *Testen und Bewerten fremdsprachlicher Kompetenzen*. Tübingen: Narr.

Dieses Studienbuch wendet sich sowohl an Studierende als auch PraktikerInnen und gibt einen guten Einblick in die relevanten Aspekte des Testens und Bewertens. Dabei werden der aktuelle Forschungsstand und praktische Anwendungsbeispiele mit einbezogen.

Weiterführende Literatur zu Testgrundlagen / Entwicklung / Validierung

ALTE Principles of good practice (2020): [https://www.alte.org/resources/Documents/ALTE%20Principles%20of%20Good%20Practice%20Online%20\(Final\).pdf](https://www.alte.org/resources/Documents/ALTE%20Principles%20of%20Good%20Practice%20Online%20(Final).pdf)

Association of Language Testers in Europe (ALTE) Manual for Language Test Development and Examining (2011): https://www.alte.org/resources/Documents/ManualLanguageTest-Alte2011_EN.pdf

Bae, Jungok & Lee, Yae-Sheik (2011): The validation of parallel test forms: 'Mountain' and 'beach' picture series for assessment of language skills. *Language Testing* 28: 2, 155-177.

Biebricher, Christine (2008): *Lesen in der Fremdsprache. Eine Studie zu Effekten extensiven Lesens*. Tübingen: Narr.

Bortz, Jürgen & Döring, Nicola (2002): *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. überarb. Aufl.). Berlin: Springer.

Cheng, Liying; Watanabe, Yoshinori & Curtis, Andy (2004): *Washback in Language Testing: Research Contexts and Methods*. Mahwah, NJ: Erlbaum.

- Cizek, Gregory J. & Bunch, Michael B. (2007): *Standard-setting. A guide to establishing and evaluating performance standards on tests*. California: Sage.
- Council of Europe (2009): *Relating language examinations to the Common European Framework of Reference for Languages (CEF). A Manual*. Straßburg: Language Policy Division.
- Council of Europe Grids to specify language tests (n.d.): <https://www.coe.int/en/web/common-european-framework-reference-languages/relating-examinations-to-the-cefr>
- Cronbach, Lee J. (1988): Five Perspectives on Validity Argument. In: Wainer, Howard & Braun, Henry (Hrsg.): *Test Validity*. Hillsdale, NJ: Lawrence Erlbaum, 3-17.
- Dlaska, Andrea & Krekeler, Christoph (2009): *Sprachtests. Leistungsbeurteilungen im Fremdsprachenunterricht evaluieren und verbessern*. Hohengehren: Schneider.
- Dutch Grid: <https://www.lancaster.ac.uk/fss/projects/grid/>
- EALTA Guidelines (2006): <http://www.ealta.eu.org/guidelines.htm>
- Eckes, Thomas (2011): *Introduction to Many-Facet Rasch Measurement. Analyzing and Evaluating Rater-Mediated Assessments*. Frankfurt a. M.: Lang.
- Fulcher, Glenn & Davidson, Fred (2007): *Language Testing and Assessment: An Advanced Resource Book*. London and New York: Routledge.
- Fulcher, Glenn (2010): *Practical Language Testing*. London: Hodder Education.
- Green, Anthony (2007): Washback to learning outcomes: a comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy & Practice* 14, 75-97.
- Grotjahn, Rüdiger (2007): Qualitätsentwicklung und -sicherung: Gütekriterien von Tests und Testaufgaben. *Der Fremdsprachliche Unterricht Französisch* 88, 44-46.
- Grotjahn, Rüdiger (2010): Sprachtests: Formen und Funktionen. In: Hallet, Wolfgang & Königs, Frank G. (Hrsg.): *Handbuch Fremdsprachendidaktik*. Seelze-Velber: Kallmeyer, 211-215.
- Harsch, Claudia (2012): Der Einsatz von Sprachtests in der Fremdsprachenforschung: Tests als Untersuchungsgegenstand und Forschungsinstrument. In: Doff, Sabine (Hrsg.): *Fremdsprachenunterricht empirisch erforschen: Grundlagen, Methoden, Anwendung*. Tübingen: Narr, 150-183.
- Harsch, Claudia (2016, Neuauflage im Druck): Testen. In: Daniela Caspari, Friederike Klippel, Michael Legutke & Karen Schramm (Hrsg.): *Forschungsmethoden in der Fremdsprachendidaktik. Ein Handbuch*. Tübingen: Narr, 204-217.
- Harsch, Claudia (2019): What it means to be at a CEFR level. Or why my Mojito is not your Mojito – on the significance of sharing Mojito recipes. In: Huhta, Ari; Figueras, Neus & Erickson, Gudrun (Hrsg.): *Developments in language education – a memorial volume in honour of Sauli Takala*. EALTA / University of Jyväskylä, Finland, 76-93. <http://www.ealta.eu.org/documents/resources/Developments%20in%20Language%20Education%20A%20Memorial%20Volume%20in%20Honour%20of%20Sauli%20Takala.pdf>.
- International Language Testing Association (2018): Code of Ethics. <https://www.iltaonline.com/page/CodeofEthics>.
- International Language Testing Association (2020): ILTA guidelines for practice. <https://www.iltaonline.com/general/custom.asp?page=ILTAGuidelinesforPractice>.
- Kane, Michael T. (1992): An argument-based approach to validity. *Psychological Bulletin* 112: 3, 527-535.
- Kane, Michael T. (2001): Current Concerns in Validity Theory. *Journal of Educational Measurement* 38, 31-342.
- Lienert, Gustav & Raatz, Ulrich (1998): *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz.
- Linacre, John M. (1994): *Many-Facet Rasch Measurement* (2. Aufl.). Chicago: MESA Press.
- Marx, Nicole (2005): *Hörverstehensleistungen im Deutschen als Tertiärsprache. Zum Nutzen eines Sensibilisierungsunterrichts im „DaF-nE“*. Baltmannsweiler: Schneider Hohengehren.
- McNamara, Tim & Roever, Carsten (2006): *Language testing: The social dimension*. London: Blackwell.
- McNamara, Tim (2006): Validity in language testing: The challenge of Sam Messick's legacy. *Language Assessment Quarterly* 3: 1, 31-51.

- Messick, Samuel A. (1989): Validity. In: Linn, Robert L. (Hrsg.): *Educational Measurement*. (3. Aufl.). Washington DC: The American Council on Education and the National Council on Measurement in Education, 13-103.
- Messick, Samuel A. (1994): The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 23: 2, 12-23.
- Mislevy, Robert J. (2007): Validity by Design. *Educational Researcher* 36: 8, 463-469.
- Rossa, Henning (2012): *Mentale Prozesse beim Hörverstehen in der Fremdsprache. Eine Studie zur Validität der Messung sprachlicher Kompetenzen*. Frankfurt a. M.: Lang.
- Taylor, Lynda & Weir, Cyril J. (Hrsg.) (2009): *Language Testing Matters: Investigating the wider social and educational impact of assessment – Proceedings of the ALTE Cambridge Conference, April 2008*. Cambridge: Cambridge ESOL and Cambridge University Press.
- Taylor, Lynda (2009): Developing Assessment Literacy. *Annual Review of Applied Linguistics* 29, 21–36.
- Wall, Dianne (2005): *The Impact of High-Stakes Examinations on Classroom Teaching: A Case Study Using Insights from Testing and Innovation Theory*. Cambridge: Cambridge ESOL and Cambridge University Press.
- Weir, Cyril (2005): *Language Testing and Validation*. Oxford: Palgrave.

Ausgewählte Large-scale Assessment-Studien

- Beck, Bärbel & Klieme, Eckhardt (Hrsg.) (2007): *Sprachliche Kompetenzen – Konzepte und Messungen. DESI Studie*. Weinheim: Beltz.
- DESI-Konsortium (Hrsg.) (2008): *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Harsch, Claudia, Pant, Hans Anand & Köller, Olaf (Hrsg.) (2010): *Calibrating Standards-based Assessment Tasks for English as a First Foreign Language. Standard-setting Procedures in Germany*. Münster: Waxmann.
- Rupp, André; Vock, Miriam; Harsch, Claudia & Köller, Olaf (2008): *Developing Standardsbased Assessment Tasks for English as a First Foreign Language – Context, Processes and Outcomes in Germany*. Münster: Waxmann.