

Die Niveaustufen-Zuordnung des TestDaF zum Gemeinsamen europäischen Referenzrahmen für Sprachen (GER)

Zielsetzung

Das TestDaF-Institut hat in den Jahren 2005 bis 2008 an dem Projekt des Europarats „Relating Language Examinations to the Common European Framework“ teilgenommen. Zielsetzung dieses Projekts war die Erprobung eines Verfahrens für die Zuordnung von Sprachprüfungen zum GER. Das in einem Handbuch (dem Manual) dargestellte Verfahren der Niveaustufen-Zuordnung konnte auf der Grundlage von Pilotierungsstudien optimiert werden. Eine dieser Studien befasste sich mit dem TestDaF. Der Europarat hat 2009 eine Endfassung des Manuals vorgelegt.¹

Hintergrund

Mit dem Projekt verfolgte der Europarat eine Sprachpolitik, die darauf ausgerichtet ist, innerhalb des europäischen Sprachraums Transparenz und Vergleichbarkeit von Sprachtests und Sprachprüfungen herzustellen. Als Bezugsrahmen diente dabei der GER, dessen Niveaustufenskalen und Niveaubeschreibungen weite Verbreitung unter Anbietern von Sprachtests und Sprachkursen sowie Lehrbuchverlagen gefunden hat. Der Anspruch dieser Interessengruppe, Lehrbücher, Sprachprüfungen oder -kurse anzubieten, die dem Kompetenzniveau einer bestimmten Niveaustufe des GER entsprechen, sollte durch das o.g. Projekt auf eine nachvollziehbare empirische Basis gestellt werden. Denn in seinen Empfehlungen an die Mitgliedsstaaten hat z. B. das Ministerkomitee des Europarats 2008 die Befürchtung geäußert, dass aus kommerziellem Interesse der Bezug von solchen Produkten zum GER in Wirklichkeit nicht dem zugrunde liegenden Konzept des GER entspricht. Er hat dementsprechend folgende Empfehlung ausgesprochen, die die Transparenz und Qualitätssicherung bei der Verwendung des GER garantieren soll:

„[...] ensure that all tests, examinations and assessment procedures leading to officially recognised language qualifications take full account of the relevant aspects of language use and language competences as set out in the CEFR, that they are conducted in accordance with internationally recognised principles of good practice and quality management, and that the procedures to relate these tests and examinations to the common reference levels (A1-C2) of the CEFR are carried out in a reliable and transparent manner;“ (Council of Europe 2008: 4)²

Untersuchungsdesign

Entsprechend der vom Europarat angeregten Validierung von Niveaustufen-Zuordnungen verfolgte die von Gabriele Kecker geleitete TestDaF-GER-Studie das Ziel, die Zuordnung des TestDaF zu den Niveaustufen B2 und C1 des GER empirisch zu überprüfen.

Das Untersuchungsdesign berücksichtigte alle vier methodischen Phasen wie sie im Manual dargelegt sind: Testspezifikation anhand der GER-Skalen, Nachweis der internen Validität des

¹ Council of Europe. (2009). *Relating language examinations to the common European framework of reference for languages: Learning, teaching, assessment (CEFR): A manual*. Strasbourg: Council of Europe/Language Policy Division. Retrieved from http://www.coe.int/t/dg4/linguistic/Manuel11_EN.asp.

² Council of Europe (2008). *Recommendation CM/Rec(2008)7 of the Committee of Ministers to member states on the use of the Council of Europe's Common European Framework of Reference for Languages (CEFR) and the promotion of plurilingualism*. Strasbourg: Committee of Ministers.

TestDaF, Standard-Setting und empirische Validierung. Bei der Durchführung der einzelnen Schritte wurden sowohl qualitative als auch quantitative Untersuchungsmethoden angewendet.

Für die Phase des Standard-Settings wurde in den rezeptiven Teilkompetenzen die Basket-Methode ausgewählt und in den produktiven die Benchmarking-Methode. Beim Standard-Setting handelt es sich um ein Verfahren, in dem mithilfe von Expertenurteilen und/oder statistischen Methoden Bestehensgrenzen für Kompetenzstufen (Cut-Offs) festgelegt werden. Die Basket-Methode arbeitet mit der Vorstellung eines kompetenten Testkandidaten auf einer bestimmten GER-Stufe und erfordert von den Experten die Antwort auf die Frage: „Auf welcher Niveaustufe kann ein Testteilnehmer das folgende Item schon richtig beantworten?“. Die Anzahl der Items, die jeder einzelne Experte einer bestimmten GER-Niveaustufe zuordnet, wird ermittelt und der Durchschnitt über alle Experten gebildet. Das Ergebnis wird abgerundet und bildet den Cut-Off. Bei der Benchmarking-Methode werden von Experten standardisierte mit Bezug zum GER kalibrierte Leistungsbeispiele als Vergleichsgröße für die Zuordnung der zu bewertenden Leistungen verwendet, um die relevanten GER-Stufen im schriftlichen und mündlichen Ausdruck festzulegen. In drei Workshops mit insgesamt 41 Experten wurden nach diesen Methoden Bestehensgrenzen für die GER-Niveaustufen B2 und C1 im TestDaF ermittelt.

Die Phase der empirischen Validierung diente der erneuten Überprüfung der GER-Zuordnung der TestDaF-Niveaustufen mit zwei Außenkriterien: Lehrerurteilen zur Sprachkompetenz von TestDaF-Teilnehmern aus Vorbereitungskursen und Ergebnisse von TestDaF-Teilnehmern im DIALANG. In beiden Fällen wurden die anhand der Außenkriterien ermittelten GER-Niveaustufen mit den im TestDaF erzielten Ergebnissen und deren Entsprechung im Stufensystem des GER korreliert.

Ergebnisse

Die Anwendung dieses Verfahrens auf den TestDaF hat eine weitgehende Übereinstimmung der bisherigen Zuordnung zu den GER-Niveaustufen B2 und C1 mit der empirisch ermittelten ergeben. Vereinzelt Abweichungen lassen sich z. B. auf die Schwierigkeit zurückführen, die zum Teil wenig kohärenten und allgemein gehaltenen Niveaustufen-Beschreibungen des GER, die Sprachverarbeitungsprozesse häufig nur abstrakt angeben, mit den Testaufgaben einer Sprachprüfung zu verbinden, deren schwierigkeitsgenerierende Merkmale sehr gezielt auf solche Prozesse abgestimmt sind. Eine ausführlichere Darstellung der Untersuchung und ihrer Ergebnisse findet sich in Kecker und Eckes (in Druck).

Dr. Gabriele Kecker

Hagen, 01.09.2010