

Takala, Sauli

Relating Examinations to the Common European Framework

Klieme, Eckhard [Hrsg.]; Beck, Bärbel [Hrsg.]: Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim u.a. : Beltz 2007, S. 306-313

urn:nbn:de:0111-opus-31485

in Kooperation mit:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

pedocs gewährt ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit dem Gebrauch von pedocs und der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt:

peDOCS

Deutsches Institut für Internationale Pädagogische Forschung (DIPF)

Informationszentrum (IZ) Bildung

Schloßstr. 29, D-60486 Frankfurt am Main

eMail: pedocs@dipf.de

Internet: www.pedocs.de

Bärbel Beck / Eckhard Klieme (Hrsg.)

Sprachliche Kompetenzen

Konzepte und Messung

DESI-Studie

(Deutsch Englisch Schülerleistungen International)

Beltz Verlag · Weinheim und Basel

Dr. *Bärbel Beck* ist Diplompsychologin und Projektkoordinatorin am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt a.M.

Prof. Dr. *Eckhard Klieme* ist Direktor des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) in Frankfurt a.M.

Diese Studie wurde im Auftrag der Kultusministerkonferenz erstellt. Für die Richtigkeit des Ergebnisses der Studie trägt das »Deutsche Institut für Internationale Pädagogische Forschung« allein die Verantwortung.

Das Werk und seine Teile sind urheberrechtlich geschützt. Jede Nutzung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages. Hinweis zu § 52a UrhG: Weder das Werk noch seine Teile dürfen ohne eine solche Einwilligung eingescannt und in ein Netzwerk eingestellt werden. Dies gilt auch für Intranets von Schulen und sonstigen Bildungseinrichtungen.

Lektorat: Peter E. Kalb

© 2007 Beltz Verlag · Weinheim und Basel

www.beltz.de

Herstellung: Klaus Kaltenberg

Satz: Deutsches Institut für Internationale Pädagogische Forschung

Druck: Druckhaus »Thomas Müntzer«, Bad Langensalza

Printed in Germany

ISBN 978-3-407-25398-9

Inhaltsverzeichnis

<i>Bärbel Beck / Eckhard Klieme</i> Einleitung.....	1
--	---

Übergreifende Konzeptualisierung sprachlicher Kompetenzen

<i>Nina Jude / Eckhard Klieme</i> Sprachliche Kompetenz aus Sicht der pädagogisch-psychologischen Diagnostik.....	9
---	---

<i>Günter Nold / Heiner Willenberg</i> Lesefähigkeit	23
---	----

<i>Claudia Harsch / Astrid Neumann / Rainer Lehmann / Konrad Schröder</i> Schreibfähigkeit.....	42
--	----

<i>Wolfgang Eichler / Günter Nold</i> Sprachbewusstheit	63
--	----

Messung sprachlicher Kompetenzen

<i>Johannes Hartig</i> Skalierung und Definition von Kompetenzniveaus	83
--	----

<i>Jürgen Rost</i> Definition von Kompetenzniveaus mit Hilfe von Mischverteilungsmodellen	100
---	-----

Kompetenzmodelle und Kompetenzniveaus im Bereich des Deutschen

<i>Heiner Willenberg</i> Lesen.....	107
--	-----

<i>Heiner Willenberg / Steffen Gailberger / Michael Krelle</i> Argumentation	118
---	-----

<i>Heiner Willenberg</i> Wortschatz.....	130
---	-----

<i>Günther Thomé / Jens Gomolka</i> Rechtschreiben.....	140
--	-----

<i>Wolfgang Eichler</i> Sprachbewusstheit	147
--	-----

Albert Bremerich-Vos / Rüdiger Grotjahn
Lesekompetenz und Sprachbewusstheit:
Anmerkungen zu zwei aktuellen Debatten 158

Kompetenzmodelle und Kompetenzniveaus im Bereich des Englischen

Günter Nold / Henning Rossa
Hörverstehen 178

Günter Nold / Henning Rossa
Leseverstehen 197

Claudia Harsch / Konrad Schröder
Textrekonstruktion: C-Test 212

Günter Nold / Henning Rossa
Sprachbewusstheit 226

Günter Nold / John H. A. L. De Jong
Sprechen 245

Hermann-Günter Hesse / Kerstin Göbel
Interkulturelle Kompetenz 256

Günther Schneider
Auf dem Weg zu Skalen für die rezeptiven
Kompetenzen im Bereich des Englischen 273

Ausblick

Konrad Schröder
Kompetenz, Bildungsstandards und Lehrerbildung
aus fachdidaktischer Sicht 290

Günter Nold
DESI im Kontext des Gemeinsamen
Europäischen Referenzrahmens für Sprachen 299

Sauli Takala
Relating Examinations to the Common European Framework 306

Hermann Lange
Abschließendes Statement 314

Die Autorinnen und Autoren 318

Sauli Takala

Relating Examinations to the Common European Framework

Introduction

A cornerstone in the domain of language learning, teaching and testing in Europe is undoubtedly the Common European Framework of Reference for Languages: Learning, teaching, assessment (Council of Europe 2001), henceforth referred to as “the CEF”. The CEF has been the central theme of many discussions and development work among test providers and individual language testers throughout Europe. Such discussions have contributed to getting a better grasp of a number of existing problems and facilitated movement towards transparency and convergence in language testing, examinations and certification.

The CEF draws on fifty years of cultural co-operation under the aegis of the Council of Europe. After a number of medium-term projects, a symposium in Rüschtikon in November 1991 produced a recommendation for developing a European Framework of Reference for language teaching and learning and for a study to be made of the feasibility of a European Language Portfolio (ELP) for reporting achievement in language skills in relation to that Framework.

Several internal drafts of the CEF were produced in 1995-6 and a first version was circulated in March 1997 for a wide-scale round of consultation before being submitted to the final conference of the project in April 1997. Feedback from piloting in the years that followed led to the publication of the Framework in English, German, French in 2001 to mark the European Year of Languages. (For a detailed account of development work related to the scales, see North 2000; North/Schneider 1998; Schneider/North 2000; a set of case studies in using the CEF is available in Alderson 2002).

The CEF as framework in testing and assessment in Europe

There has been a strongly growing interest in Europe (and indeed increasingly elsewhere) in linking examinations to the CEF. While the CEF is designed to serve as a tool for those involved in the development of syllabuses, curriculum guidelines, textbooks, teacher education, etc. (cf. p. 1), it is also specifically concerned with testing and examinations:

„One of the aims of the Framework is to help partners to describe the levels of proficiency required by existing standards, tests and examinations in order to facilitate comparisons between different systems of qualifications. For this purpose the Descriptive Scheme and the Common Reference Levels have been developed. Between them they provide a conceptual grid, which users can exploit to describe their system.“ (p. 21)

It appears that the calibrated Common Reference Levels have attracted more attention than the other key element of the CEF, the descriptive scheme. They provide a qualitatively new tool for the needs of transparency in testing and examinations. Figueras et al. (2005) have characterized the CEF as a structured collection of descriptors for a very broad range of communicative language activities and aspects of language proficiency. The structure should be understood in a qualitative and a quantitative way. Qualitatively, six main categories are distinguished in 3 x 2 setup: reception, production and interaction in the modes ‘spoken’ and ‘written’. Each of these six cells are then further subdivided. In the 2001 edition, there are 54 sets of such descriptor scales. Quantitatively, the categories for communicative language activities, strategies and aspects of language proficiency have been defined at a series of 6 levels A1, A2, (Basic User); B1, B2 (Independent User); C1, C2 (Proficient User). The CEF suggests (pp. 31-33) that more specific sub-levels can and should be distinguished if the assessment purpose and context requires this.

Since the Council of Europe Symposium in Rüşchlikon, in 1991, which launched the process that led to the CEF and the ELP, there has been more and more demand for coherence and transparency in language certification from Member States both within the Council of Europe and the European Union. Figueras et al. (2005) note that this means that the CEF is increasingly referred to as the obvious basis for the mutual recognition of language qualifications in Europe. Language testing and assessment in Europe is known (even without empirically documented studies) to display a wide range of traditions, institutional infrastructure, resources and professionalism. Figueras et al. (2005) suggest that:

„Therefore it is not surprising that following the publication of the CEF, there have been calls on several occasions for the Council of Europe to take a more active role in assisting examination providers in their efforts to situate their examinations within the Common European Framework, and in validating – in one form or another – language examinations that claim such linkage.“

In response to the expressed need for guidance to assist examination providers to relate their examinations to the CEF, the Council of Europe produced in September 2004 a preliminary pilot version of a Manual for Relating language Examinations to the Common European Framework of Reference (Council of Europe 2003) - referred to from now on as “the Manual” (*www.coe.int*). A reference supplement with more technical information is also available on the website.

The Manual: Procedures proposed for linking examinations to the CEF

The primary aim of the Manual is to help the providers of examinations to develop, apply and report transparent, practical procedures in order to relate their examination(s) to the CEF. The approach developed offers guidance to users to:

- describe the examination coverage, administration and analysis procedures;
- relate results reported from the examination to the “Common Reference Levels” presented in Chapter 3 of the CEF;
- provide supporting evidence that reports the procedures followed to do so.

The Manual has wider aims to actively promote and facilitate co-operation among relevant institutions and experts in member countries. The Manual aims to:

- contribute to competence building in the area of linking assessments to the CEF;
- encourage increased transparency on the part of examination providers;
- encourage the development of both formal and informal national and international networks of institutions and experts.

Relating an examination or test to the CEF is a complex endeavour. The existence of such a relation is not a simple observable fact, but is an assertion for which the examination provider needs to provide both theoretical and empirical evidence. The procedures by which such evidence is put forward can be summarized by the term “validation of the claim”. It is important to note that while the Manual covers a broad range of activities, its aim is limited:

- It provides a guide specifically focussed on procedures involved in the validation of a claim that a certain examination or test is linked to the CEF.
- It does not provide a general guide how to construct good language tests or examinations. There are several useful guides that do this and they should be consulted. Relating examinations to CEF makes sense only if the examinations are of good quality.
- It does not prescribe any single approach to constructing language tests or examinations. While the CEF espouses an action-oriented approach to language learning and use, being comprehensive it accepts that different examinations reflect various goals (“constructs”). Before embarking on relating an examination to CEF, it is the prior responsibility of the examination providers to demonstrate the validity of their examination by showing that it assesses the constructs intended.

Relating examinations to the CEF can best be seen as a process of “building an argument” based on a theoretical rationale. As noted above, the central concept within this process is “validity”, and this can be seen in a graphical representation of the process (Figure 1).

The Manual presents four inter-related sets of procedures that users are advised to follow in order to design a linking scheme in terms of a set of required activities. A fundamental prior requirement is that the examination has been shown to produce

reliable scores and yield valid interpretations. The activities carried out in all four sets of procedures contribute to the validation process.

Familiarisation: a selection of activities designed to ensure that participants in the linking process have a detailed knowledge of the CEF. This familiarisation stage is necessary at the start of both the Specification and the Standardisation procedures. In terms of validation, these procedures are an indispensable starting point. An account of the activities taken and the results obtained is an essential preliminary component of the validation report.

Specification: a self-audit of the coverage of the examination (content and tasks types) profiled in relation to the categories presented in CEF Chapter 4 “Language use and the language learner” and CEF Chapter 5 “The user/learner’s competences.” As well as serving a reporting function, this exercise also has a certain awareness-raising function that may assist in further improvement in the quality of the examination concerned. These procedures assure that the definition and production of the test have been undertaken carefully, following good practice.

Standardisation: suggested procedures to facilitate the implementation of a common understanding of the “Common Reference Levels” presented in CEF Chapter 3. Standardised exemplars will be provided to assist training in the standardisation of judgements.

These procedures assure that judgements taken in rating performances reflect the constructs described in the CEF, and that decisions about task and item difficulty are taken in a principled manner on the basis of evidence from pre-testing as well as expert judgement.

Empirical Validation: the collection and analysis of test data and ratings from assessments to provide evidence that both the examination itself and the linking to the CEF are sound. Suggestions and criteria are provided for adequate and credible validation appropriate for different contexts. These procedures assure that the claims formulated through Specification and Standardisation (“test-under-construction”) can indeed be confirmed when the examination is administered in practice (“test-in-action”) and data on how persons belonging to the target population behave when the test is so administered becomes available.

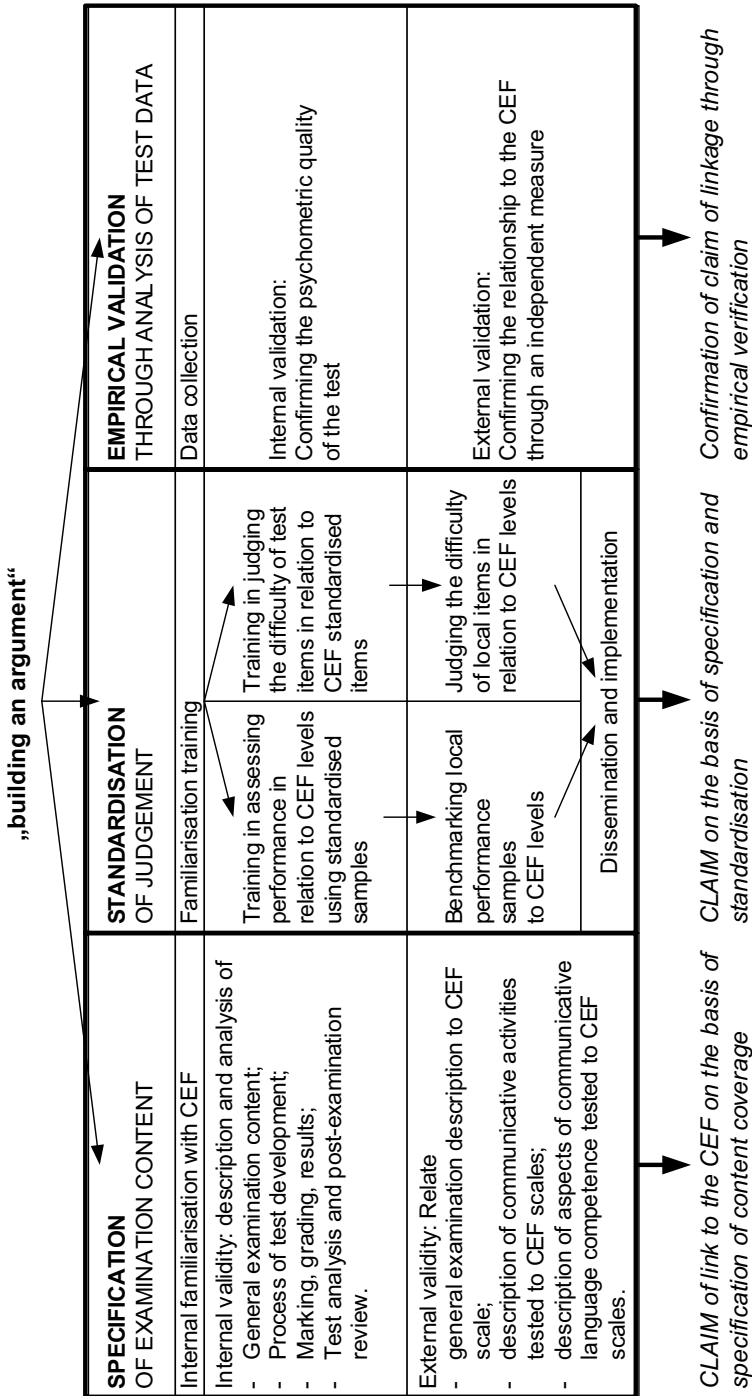


Figure 1: Visual representation of the procedures to relate examinations to the CEF: theoretical rationale (“building an argument”).

Since the Manual concerns linking examinations to the CEF, the issue of standard setting – the setting of performance standards – is central to the endeavour. Standard setting is the process of setting and validating the “cut off score(s)” for the different grades reported from a test. In the context of the Manual, this means using standard-setting procedures to allocate the cut off scores for the CEF level(s) reported by the test (in Standardisation) and then validating and if necessary refining the cut-offs through analysis of test data, preferably in connection with an external criterion (in Empirical Validation).

While standard setting in language education is relatively new in a European context, it is not a new development. In fact, systematic standard setting in education has been a topic of interest for 50 years, as the first article on this theme was published by Nedelsky in 1954. Since then dozens of methods have been proposed (for a good review of the developments in standard setting, see Cizek 2001; Kaftandjieva 2004). The growing interest in using content and performance standards and the high stakes involved are also reflected in the profession’s growing concern with the criteria for standard-setting. In the two consecutive editions of the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education 1985, 1999) the standards related to standard setting have increased from 6 to 13. The linkage between language examinations and the CEF can be established in at least three different ways:

- Direct linkage to the CEF scales of language proficiency (as discussed so far).
- Indirect linkage via linkage to a local scale of language proficiency, which has itself already been linked to the CEF scales (so far, such scales have been constructed at least in Finnish and Catalan).
- Indirect linkage via equation to an existing test that has already been linked to the CEF scales.

Irrespective of the approach adopted in the particular concrete situation, the linkage always requires some form of standard-setting. In other words, standard-setting is in the core of the linkage process (Kaftandjieva 2004).

Some issues to be addressed

In spite of the fact that the CEF offers considerably greater explicitness than most curricular documents, the CEF is still quite an abstract descriptive system. Thus to answer the following basic question is not an easy task:

By which procedures can a person be assigned to one of the described levels on the basis of his/her test performance in such a way that the assigned level corresponds to the level as meant by the descriptions in the CEF?

Another challenge is posed by the commonly put question:

“How can I be sure that your B1 (in country A for language L with test X) is my B1 (in country B for language M using test Y)?”

The questions do not stop here. Further issues to be resolved include e.g. these:

- How many items at a given level do test takers need to get right, and/or what proportion of the descriptors at a given level need to apply to a test taker for them to be assigned to a particular level?
- What standard-setting strategies would be the most appropriate in setting cut-off scores (compensatory, conjunctive, mixed/hybrid, disjunctive; Kaftandjieva 2004) and what external validation procedures are necessary to corroborate their results?
- How should scores from different sections and sub-tests be summarized in order to produce one score for each skill, and report an overall global CEF level?

A special issue of the journal of Language Testing to appear in the autumn of 2005 deals with such issues and some preliminary clarifications are provided by Figueras et al. (2005).

References

- Alderson, J.C.A. (Ed.) (2002): *Case Studies in applying the Common European Framework*. Strasbourg, Council of Europe.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1985): *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999): *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Cizek, G.J. (Ed.) (2001): *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, N.J.: Erlbaum, pp 3-17.
- Council of Europe (2001): *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge. Cambridge University Press. (www.coe.int).
- Council of Europe (2003): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual, Preliminary Version*. http://www.coe.int/T/e/Cultural_Co-operation/Education/Languages/Language_Policy/Manual/Manual.pdf
- Figueras, N./North, B./Takala, S./Verhelst, N./Van Avermmaet, P. (2005): *Relating Examinations to the Common European Framework: A Manual*. *Language Testing*, 22, 3, pp 261-279.
- Kaftandjieva, F. (2004): *Standard setting*. Section B in the Reference Supplement to the Manual (www.coe.int).
- Nedelsky, L. (1954): *Absolute grading standards for objective tests*. *Educational and Psychological Measurement* 14, pp 3-19.
- North, B. (2000): *The development of a common framework scale of language proficiency*. New York, Peter Lang.
- North, B./Schneider, G. (1998): *Scaling descriptors for language proficiency scales*. *Language Testing* 15, 2, pp 217-262.

Schneider, G./North, B. (2000): Fremdsprachen können: was heißt das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit. Nationales Forschungsprogramm 33: Wirksamkeit unserer Bildungssysteme. Chur/Zürich, Verlag Rüegger.