



Testing
& Evaluation
& Assessment

The Testing, Evaluation and Assessment Special Interest Group

Diversity, plurilingualism and their impact on language testing and assessment

IATEFL TESTING EVALUATION
AND ASSESSMENT SIG &
UNIVERSITA PER STRANIERI
DI SIENA

SIENA 22-23 November, 2013

Selected articles by the presenters edited by Judith Mader and Zeynep Urkun

ISBN: 978-1-901095-69-2

Publisher: IATEFL (TEA SIG)
www.iatefl.org

TABLE OF CONTENTS

4	Conference Program	
6	UNISTRASI	UNIVERSITÀ PER STRANIERI DIE SIENA
7	Judith Mader & Zeynep Urkun	EDITORS' NOTES
8	Kathrin Eberharter	CONFERENCE REPORT
9	Carla Bagna & Sabrina Machetti	INTERNATIONAL MOBILITY AND THE IMPACT ON LT: THE CASE OF CHINESE STUDENTS
13	Veronica Benigno	ASSESSING INTERCULTURAL COMMUNICATIVE COMPETENCE THROUGH VOCABULARY
17	Claudia Borghetti	CONSIDERATIONS ON DYNAMIC ASSESSMENT OF INTERCULTURAL COMPETENCE
21	Rudi Camerer	TESTING INTERCULTURAL COMPETENCE IN (INTERNATIONAL) ENGLISH
26	Carol Spöttl & Kathrin Eberharter	USING THE CEFR IN DEVELOPING AND STANDARD SETTING WRITING TASKS ACROSS FOUR LANGUAGES
29	Judith Hanks	PROFESSIONAL LANGUAGE INTERCULTURAL STUDIES: WHAT AND HOW TO ASSESS, AND WHY?
35	Geraldine Ludbrook	INVESTIGATING A VALID TEST CONSTRUCT WITHIN AN ELF FRAMEWORK
39	Sabrina Machetti & Raymond Siebetchou	PLURILINGUAL IMMIGRANT REPERTOIRES: WHAT'S UP?

Investigating a Valid Test Construct Within an ELF Framework

Geraldine Ludbrook

Ca' Foscari University of Venice

Introduction

English is used as a Lingua Franca throughout the world, essentially as a means of communication between speakers of different first languages. In academic settings, English has been used for decades in teaching and learning, as well as in intercultural communication. English as a medium for education has been in constant expansion in continental Europe since the 1999 Bologna Agreement, and the subsequent increase in Erasmus and similar exchange student agreements, and English-taught degree programmes.

The high-stakes international tests adopted to assess the language of students seeking to study in English-medium universities are generally developed with reference to the standard English of the countries in which they are to be applied. The same tests are used to measure the English language proficiency of European students applying for admission to English-medium university courses in contexts in which English is used as a Lingua Franca between university teachers and students, among students, and between administrators and students. In these settings, in which the focus is on effective communication rather than accuracy, native speaker competence may no longer be an appropriate benchmark.

Advocates of English as a Lingua Franca (ELF) have long called for international examination boards to consider basing the criteria used in tests on successful communication between non-native speakers of English rather than between native speakers, or between a non-native speaker and a native speaker. Examination boards have to a degree moved in this direction: in recent years a focus on knowledge and form has given way to a greater emphasis on function and communication, and communicative effectiveness has replaced adherence to standard norms. However, Standard English, as it is codified in grammars and dictionaries, does appear to remain the default model for language testing.

This paper briefly reports on a research project aimed at exploring some of the issues involved in the development of a computer-delivered language test assessing receptive skills of reading and listening at CEFR B1/B2 level within an ELF framework, with particular attention to the definition of the construct of the test.

Towards a valid test construct

There are many challenges entailed in developing a test within an ELF framework which have been examined in the testing literature (see, for example, Elder and Harding, 2008). A central factor to the validity of a test lies in the validity of the construct, the language abilities to be measured in the test, or the fitness of the test for its purpose and context. For test developers, therefore, having as clear and unambiguous a definition as possible of what the construct should include is necessary at the outset of any test development project.

The challenge with a construct within an ELF context is that ELF norms are “unstable, negotiated online, continually shifting and changing, and that in the variable process of language use people do not just conform but adapt their linguistic and cultural behaviour as appropriate to their communicative purposes” (Seidlhofer, 2008). Test developers therefore must address the issue of defining a valid construct with a language model that embraces this variation and instability.

To seek a solution, the research project turned to Douglas's (2000) framework of specific purpose language ability. Douglas describes communicative language ability through two interacting components: language knowledge and strategic competence (p. 33). In Douglas's framework, *language knowledge* consists of four main features:

1. Grammatical knowledge: vocabulary, morphology, syntax, phonology,
2. Textual knowledge: rhetorical organization, cohesion,
3. Functional knowledge: ideational, manipulative, heuristic, imaginative functions,
4. Sociolinguistic knowledge: sensitivity to dialects, registers, naturalness, cultural references (p. 35).

To these features, Douglas adds four features of *strategic competence*:

1. Assessment: evaluating the communicative situation and engaging a discourse domain, a

cognitive interpretation of the context,

2. Goal setting: deciding whether and how to respond to the situation,
3. Planning: deciding what elements of language and background knowledge are required,
4. Control of execution: organizing the required elements to carry out the plan (p. 35).

Douglas also adds a third dimension to his model of specific-purpose language ability – *specific-purpose background knowledge* – asserting that specific-purpose language tests require test takers to engage authentically in test tasks that are demonstrably related to the target language use situation. Hence background knowledge plays an essential role in tests of language for specific purposes (p. 39).

Identifying target language use

Douglas conceives of a special purpose language test as one in which test content and methods are derived from an analysis of a specific purpose target language use situation, so that test tasks and content are authentically representative of tasks in the target situation, allowing for an interaction between the test taker's language ability and specific purpose content knowledge, on the one hand, and the test tasks on the other. Such a test allows us to make inferences about a test taker's capacity to use language in the specific purpose domain. The project set out to draw up a detailed picture of the use of English made by university students studying in English either in their home or in a foreign university.

The primary data for the analysis was gathered from three different sources. The first was Venice International University, where the working language for academic teaching, administration, and peer interaction is English. The methods used included semi-structured interviews with teachers and students regarding their uses of English and their perceptions of their language needs; observation of classes conducted in English by native and non-native speaking university lecturers; and recording and analysis of ELF peer interaction among students.

The second was two universities in different Italian regions. An investigation was carried out of the English language needs in the two contexts as perceived by Italian faculty and students. The data was collected using student questionnaires regarding perceived English language needs, and interviews with teachers re-

garding the skills and level of English competence needed by their students.

The third step was the examination of the websites of European universities, with particular attention to the sections providing advice to incoming and outgoing students involved in Erasmus exchange programmes. A survey of the information published on these websites provided valuable insights into the real-world contexts of undergraduate life in which foreign students are required to engage.

Analysis of the TLU domain

The analysis of the TLU domain identified three distinct but interrelated language contexts.

1. Academic language

Although the variability of norms in ELF is considerable, the norms regarding academic English are relatively stable as the domain implies the language of educated users in which the features of accuracy and appropriateness are desirable.

The main English academic language needs concerning receptive skills in the contexts of Italian students studying in home universities, and Italian students studying abroad, as perceived both by the students themselves and by faculty, were identified as:

- Reading textbooks and scholarly articles
- Using the Internet for research
- Watching film and video
- Understanding native and non-native speaking teachers
- Understanding native and non-native speaking students

2. Administrative language

The purpose of the test was to assess the English language skills of students intending to study in English-medium academic courses. The test therefore needed to include the language required to deal with the planning stage before departure. A review of the advice given on university websites to students planning to study in other European universities identified a series of areas in which students are required to retrieve practical information regarding administrative procedures.

The main English needs concerning university administration procedures were identified as:

- Understanding university application and enrolment procedures
- Understanding tuition fees and methods of

payment

- Understanding applications for funding and student loans
- Understanding study programmes, courses and examinations
- Dealing with academic problems, such as failing exams, plagiarism
- Dealing with coursework requirements, such as submission of work, deadlines

3. Everyday language

Advice given on university websites to students planning to study in other European universities also permitted the identification of a series of areas in which students can retrieve practical information regarding everyday student life abroad.

The main English needs concerning everyday university life were identified as:

- Selecting accommodation: on campus residence, student housing, host family
- Using university facilities: libraries, canteen, gym, computer laboratories, wireless networks

In conclusion, the analysis of the target language use provided a detailed picture of the language needs of European university students planning to study in courses taught through English in foreign university contexts.

Input texts

The test project was limited to the receptive skills of reading and listening/viewing. The analysis of the target language use provided detailed information regarding the characteristics of the input reading and listening texts.

The following reading input texts were identified:

- University websites
- Titles of university courses, books, lectures
- Short descriptions of university courses
- Internet search engine results
- Short general academic texts

The following audio and visual listening texts were identified:

- Presentation of general academic topics
- Explanations of academic and organizational procedures
- Discussions of topics related to everyday university

Test tasks

The analysis of the target language use domain also led to the identification of input texts and related language functions, which were then

operationalized in test tasks. In this phase, importance was given to Messick's (1994) threat to test validity that may arise when the construct of a test is operationalized in the test tasks: construct-irrelevant variance, which occurs when a test measures variables that are unrelated to the construct. This may result in construct irrelevant easiness, when the test task may provide clues that allow some test takers to respond correctly in ways that are irrelevant to the test construct, and result in higher scores. It may also result in construct irrelevant difficulty, when the test is more difficult for some test takers, for irrelevant reasons, and results in lower scores (pp. 34-35).

The test was designed to assess the English language competence of European students planning to enrol at university, not of students already enrolled in university courses. Particular care was therefore taken when designing the tasks, and choosing the input texts, to include only general academic content and to avoid introducing specific subject-related knowledge.

In addition, background knowledge concerning only common university procedures and predictable aspects of student life was included. In this way it was hoped to avoid the threat of construct-irrelevance variance, which would have weakened the validity of the test.

The test tasks were identified as:

- Reading a university website to retrieve information
- Reading titles of university courses, books, lectures to identify their content
- Reading short descriptions university courses to identify their content
- Reading an internet search engine result to retrieve information
- Reading short general academic texts to retrieve, analyse and process information
- Listening to teachers presenting general academic topics to retrieve information
- Listening to teachers and administrators explaining academic and organizational procedures to retrieve information
- Listening to peers discussing topics related to everyday university life to retrieve information

The items developed were trialed with students from the University language centre, demonstrated that many of the items developed were useful for the purpose of discriminating between the desired levels of ability B1 and B2.

Conclusions

This paper briefly reports on investigating language test design within an ELF framework. Further details on the whole project can be found in Ludbrook and Newbold (2012). The project includes features of ELF interaction and has added evidence to the body of research on ELF testing. In particular, the detailed analysis of the target language use carried out for the definition of the test construct has begun to work towards a valid construct for a test of the English ability of European university students studying and living in an academic environment in which English is used as a lingua franca.

In addition, the project has investigated the design of new kinds of test tasks that achieve high degrees of authenticity in a test developed within a specific target use domain. Situational authenticity is achieved by the close match between the test tasks and the real world tasks carried out by the test takers. Interactional authenticity is achieved by the design of test tasks based on the graphics and navigation of university internet websites and internet search engines. The high authenticity realized adds to the validity claim of the construct.

Further research would be required to add the productive skills of writing and speaking, and the attendant issues of assessment criteria. Future research would also be necessary to further investigate the other issues involved in ELF test design: questions of fairness and bias when using non-native varieties of English, and the accountability and prestige of a test developed within the ELF framework. However, we believe this initial phase of an ELF test development project is a principled basis to this future research.

References

- Douglas, D. (2000) *Language testing for specific purposes*. Cambridge: Cambridge University Press
- Elder, C. and Harding, L. (2008) Language testing and English as an international language: Constraints and contributions. In *Australian Review of Applied Linguistics*. International Forum on English as an International Language. F. Sharifian and M. Clyne (eds.) 31(3) pp. 34.1-34.11
- Ludbrook, G. and Newbold, D. (eds.) (2012) *English Lingua Franca: Contexts, Strategies and International Relations*. Venezia: Libreria Editrice Cafoscarina pp. 111-124
- Messick, S. (1994) The interplay of evidence and consequences in the validation of performance assessments. In *Educational Researcher* 23(2) pp. 13-23

- Seidlhofer, B. (2008) Of norms and mindsets. In *Australian Review of Applied Linguistics*. International Forum on English as an International Language. F. Sharifian and M. Clyne, (eds.) 31(3) pp. 33.1-33.7

Dr Geraldine Ludbrook is a researcher and lecturer in English Language and Linguistics at the Ca' Foscari University of Venice. She holds a PhD in Applied Linguistics from Lancaster University (UK). She has taken part in numerous national and international testing research projects. Her recent research focuses on language testing within an English as a Lingua Franca framework; CLIL teacher language and fair and valid language testing for students with dyslexia and related specific learning difficulties.
ludbrook@unive.it