4 Large-scale Testing of Proficiency in English: Back to Multiple Choice?

Irina Argüelles Álvarez, Iciar de Pablo Lerchundi, Rafael Herradón Díez, and José Manuel Baños Expósito
Universidad Politécnica de Madrid
irina@euitt.upm.es

Introduction

The Universidad Politécnica de Madrid UPM has recently included a university-wide compulsory subject in the new curriculum aimed at preparing students for international academic and professional situations "English for Professional and Academic Communication". Since the University requires the students to certify a B2 level as described by the Common European Framework of Reference for Languages (CEFRL) in order for them to have the right to enroll in that compulsory subject, the development process and the results of a specific test designed to this aim is the focus of this paper. The final purpose of the study is to demonstrate the validity of this multiple choice test to situate a student above or below a B2 proficiency level according to the CEFRL.

Test Specifications

Our specific context, the number of students who must certify a B2 level to enroll in the subject "English for Professional and Academic Communication" and the heterogeneous background of the more than sixty teachers in the Department of Linguistics lead us to opt for an automatic correction test made up of multiple choice- type questions. Although for the same practical reasons of application, the test does not include a listening comprehension section or an interview, previous studies (Argüelles et al. 2010) let us establish an initial hypothesis which presumes that most Spanish students at university who demonstrate a high level of proficiency in these aspects of language would obtain similar proficiency results in direct listening and speaking tests. The limitations of these types of tests are understood and assumed for practical reasons of application, and the extent to which they affect the students' results will be presented and analyzed in the following sections.

The layout of the test shows two differentiated parts although the test is not explicitly divided into those parts. The first part that evaluates aspects of grammar, consists of 65 individual items followed by the four options a, b, c and d. The second part of the test consists of three texts each worth 10 or 15 points, for a total of 35 points, and is centered on aspects more related to use of language, vocabulary and reading comprehension. The activities are adapted from a corpus of texts and tasks selected from general English course books which have been correlated to the CEFRL, covering B2 and going towards a C1 level, into a multiple choice format with four options.

Results

A total of 240 incoming students at School of Telecommunications took the pilot test. From the 240 students taking the test, 36 did not finish it, therefore, a total of 214 tests are taken into consideration for the statistical analysis. In what follows, a summary of the results of the test concerning its reliability and validity is presented.

The test layout is in two well differentiated parts. As such, it is important to analyze to what extent these two parts show similar results. The Pearson correlation coefficient between the scores obtained in part one and those in part two of the test is analyzed by means of SPSS and the result is 0.81, which is statistically significant (see Table 4.1). This means 81% concordance between the two parts of the test (Brown 1988: 98,99).

	Part two (use of language, vocabulary and reading)
Part one (grammar)	0.813 ($p<0.01$)

Table 4.1: Pearson Correlation between parts one and two of the English Proficiency Test

As for establishing validity in our test (Brown 1988: 102-105), a statistically representative sample from the 214 students who completed the multiple choice test was selected for the interview. The correlation between the results in the interview and the test is analyzed by means of SPSS and a validity coefficient of 0.83 is obtained. This correlation is statistically significant.

	English Proficiency Test
Oral Interview	0.825 ($p<0.01$)

Table 4.2: Pearson Correlation between Oral interview and English Proficiency Test

From the experiment, the results show that 29% of the 31 students interviewed reached the B2 level according to the CEFRL, whereas 71% did not reach this level of proficiency. Taking into account these percentages obtained from the oral interview and correlating them with the scores in the multiple choice proficiency tests, the cut-off score could be established in the range from 6.9 to 7.1.

Discussion and conclusions

It is understood here that this proposal could be highly unpopular when assessment and evaluation in contexts of higher education tend to introduce direct techniques and task-based and other more formative approaches, so there are some important points to clarify at this point. First, the proposal is made for a specific context where a B2 level must be proved on the part of the students enrolling in a course of professional and academic English. Second, multiple choice tests are a good alternative for increasing reliability, and the design and development of the test were conceived to maintain good levels of validity. Third, it is not intended here to suggest that an indirect test of this type can in any case substitute direct tests of different skills but rather to present it as a practical tool where other alternatives are difficult or impossible to carry out.

To summarize, an examination of the statistical results presented in this paper shows high levels of reliability and validity of this test to measure what it claims to measure in order to situate a student above or below the given B2 level of proficiency. Based on the initial pilot study, it is presumed that students who demonstrate this level of proficiency would eventually pass a competency type test of the same level as the one assessed here in a high percentage of the cases.

References

I. Argüelles, E. Martín, R. Herradón, G. Balabasquer & C. Ortiz. (2010). Análisis y descripción de la variación en el nivel de competencia en lengua inglesa de los estudiantes que ingresan en los nuevos grados de ingeniería de telecomunicación. In *AESLA*, Vigo.

J. D. Brown (1988). *Understanding research in second language learning*. Cambridge, Cambridge University Press.

Common European Framework. www.coe.int/t/dg4/linguistic/CADRE_EN.asp

Statistical Package for the Social Sciences (SPSS). www.spss.com