

Deconstructing Assessment constraints in low stakes ESP testing for tourism

Jesús García Laborda

Polytechnic University of Valencia, Spain

This paper was originally published in *Professional and Academic English: the IATEFL English for Specific Purposes SIG Journal*. Vol 16 (15th December 2006)

Summary

For a long time, much research has been devoted to the analysis, development and planning of language testing for specific purposes (Douglas, 2000; Skehan, 1984) and their communicative orientation (Davies, 2001). Unfortunately, very few teachers and instructors¹ are fully aware of the incidental issues that should account for when designing in low stake testing. In fact, although certain publications have presented practical issues in designing ESP tests (Douglas 2000), it was necessary to summarise the latest matters to make most instructors aware of the issues that should be considered in planning their tests for ESP students, especially in Travel & Tourism. In order to do so, this paper analysed the current literature in the field and, more important, tried to draw some guidelines that could be considered in the instructors' evaluation job. To better address this issue, the different factors affecting tests in ESP for Tourism have been divided into two groups: those related to the construction of the test and those which are not considered implicit to the test construction.

Keywords: Testing, ESP, Tourism.

¹ Although a distinction between teachers and instructors could be drawn, in this paper both words are used indistinctly.

Introduction

For a long time, much research has been devoted to the analysis, development and planning of language testing for specific purposes (Douglas, 2000; Skehan, 1984) and their communicative orientation (Davies, 2001). Unfortunately, very few teachers and instructors² are fully aware of the incidental issues that should account for when designing in low stake testing. In fact, although certain publications have presented practical issues in designing ESP tests (Douglas 2000), it was necessary to summarise the latest matters to make most instructors aware of the issues that should be considered in planning their tests for ESP students, especially in Travel & Tourism. In order to do so, this paper analysed the current literature in the field and, more important, tried to draw some guidelines that could be considered in the instructors' evaluation job. To better address this issue, the different factors affecting tests in ESP for Tourism have been divided into two groups: those related to the construction of the test and those which are not considered implicit to the test construction (diagram 1).

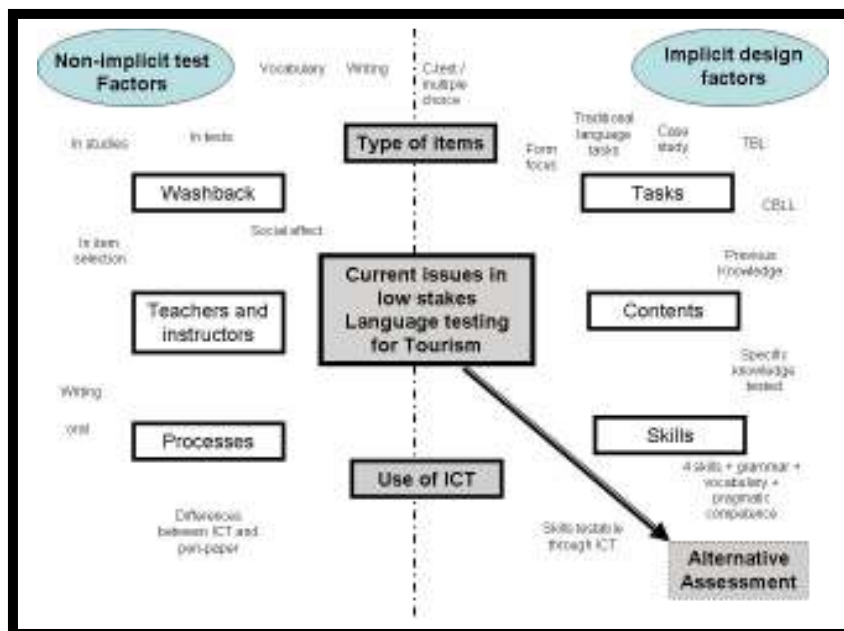


Diagram 1: Factors affecting tests in ESP for Tourism

Implicit factors in test design

² Although a distinction between teachers and instructors could be drawn, in this paper both words are used indistinctly.

Most of the circumstances and issues described in this paper are related to what is known by “low stakes tests”. To simplify, low stakes³ exams or tests are those that are distributed, rated and graded by the teacher personally. Although there is little question that the factors described in diagram 1 affect the students’ course grade, their effect is usually limited to this situation (or even less significant ones as quizzes or midterm exams). Low stakes are opposed to high stakes exams which usually are used to assess and allow students to pursue higher graduate education in different countries (TOEFL) or being accepted as professionals abroad (BULATS). Therefore, high stakes exams have more influence for a person’s future than low stakes exams. Low stakes exams are usually designed by the teachers to assess the students previous knowledge before coming to the class (diagnostic test), how the student is progressing (progress test) or how much it has been learned in the class (achievement test). Low stakes exams are usually designed by the teacher or, in many instances, included in the teacher’s textbook while high stakes ones are designed by specialists for just one use. In this article we have addressed low stakes exams because sometimes practitioner teachers are not fully aware of the different factors that can affect their assessment, especially in ESP.

Tasks and skills

One of the things that ESP instructors should consider first is whether the tests are evaluating what is taught in class. It is not unusual that teachers who emphasize the subject contents (like in Content Based instruction) may be testing grammar and reading primarily. Consequently, it is important to find the right balance between content and language use. To solve this problem, teachers may want to revise the type of items they use and what for. A possible solution is to devote compositions to content questions. Consequently it would be natural to allow students to express their ideas about sustainable tourism in the Spanish East coast in a 200 word composition. Many teachers still believe that the explicit grammar knowledge can be immediately transferred to language use but this supposition has not been supported in the last research except in writing (Bitchener, 2005). Nobody can deny that it is necessary to include new types of

³ “Low-stakes test has no significant, tangible, or direct consequences attached to the results, with information alone assumed to be a sufficient incentive for people to act” extracted from Robert Mason Hauser (Ed.) and Jay Philip Heubert (Ed.) (1999) *High Stakes: Testing for Tracking, Promotion, and Graduation*. National Research Council (U. S.)

tasks in language exams. How about including cooperative tasks or webquests in which the students need to prove their professional skills as well as their language abilities?

According to the previous point, a webquest (such as designing a trip or providing information) would show the student's efficacy in solving problems in English but there might be a problem related to the student's previous knowledge of the matter. A few studies have shown the importance of this background knowledge in solving test tasks in languages for specific purposes (Salmani-Nodoushan, 2003; Alvermann & Hynd, 1989). In general, ESP teachers should either include free-content tasks or tasks whose content should be well known by the students. Lately, some ESP teachers are beginning to cooperate with other instructors to work together contents and language. Under this collaborative premise, some content that has been addressed in another core subject could be tested in the ESP classroom. Otherwise, some students may have an initial advantage and their topic knowledge may invalidate the test results. It is obvious that teachers will have to be very careful to evaluate what has been addressed in their classroom. This is also one of the reasons why using the publishers' standard tests may be not as accurate or valid as most teachers may think. Nevertheless, these standard tests are now usually produced in ways that teachers can modify them to match better their needs and classroom contents.

In relation to the skills, it is advisable to include the traditional skills of listening, writing, reading and, whenever possible, speaking activities. Additionally, many teachers believe that tests should have grammar, specific vocabulary and pragmatic competence tasks. These might include specific registers as those found in letters, memos, professional dialogues and other tasks.

The role of ICT in testing

The use of computers in language testing has increased in the last 5 or 6 years but it has been there for a much longer time. Although Internet based tests are common today it is important to value them in higher education because they save time and funds when evaluating large quantities of students (Chapelle & Douglas, 2006). García Laborda (2006) has stressed their importance in diagnosing ESP international students. Only since some new testing platforms like PLEVALEX, the new IELTS or the iBT TOEFL

have become operative, the possibility of rapid and easy-to-distribute global exams has been implemented. Overall, testing platforms are called to have a relevant space in the future of language testing but this is especially important for ESP because they allow to create specific data banks and make them accessible to most ESP practitioners. If reading, listening and vocabulary are important in language testing, ICT's offer an unique opportunity to incorporate diagrams and, even more relevant, audiovisuals that permit to provide the students with real cues in the assessment process. These and other contextual cues such as real readings help to improve the students' performance in tests and rise the students' grades. Therefore, ICT's incorporate elements that until recently were difficult to find or unrealistic.

Non implicit factors in test design

There are other factors that may affect the testing process that do not influence the test design itself but are ultimately linked to the process of assessment. Before, it was mentioned the inclusion of different skills in the exam but some teachers like certain types of items that usually mean different types of test strategy use. Say, multiple choice tests are quite different from open question or communicative tests. One of the characteristics of test construct is to address the contents and registers studied in the classroom. Thus, it would be advisable to include oral and written strategy use for real texts and visuals. Also instructors should consider the students' needs when choosing these items but, more important, tests should be aimed to linking the learners' deficiencies and the skills developed in class and the expected goals and outcomes of the course. For example, teachers who feel that their Tourism students are weak at the beginning of the course and emphasize oral skills would probably be wrong to assign more importance to grammar, vocabulary or writing in their tests (and this is a common case!). This reflective testing also may affect two last issues: the teacher as evaluating agents and the washback process.

An Amengual Pizarro (2006) state that when teachers are trained and there is a double correction there is also a rise in quality and objectivity in the assessment process. However, low stakes are rated twice almost never and instructors are specifically trained in very few institutions and usually face classes after a very short course. As a consequence, they should be extremely careful when including the cultural contents,

matching the items with what has been studied, rating the tests (it would be advisable to do it in ways that may not be affected by their relation with each student), demanding minimal requirements (average proficiency level required from each learner) and assessing the communicative tasks. These factors may potentially change the students' final evaluation and, hence, their final grade.

Finally, the washback effect (the effect of how tests affect and change language teaching), should serve to show the teachers either what should be changed in the tests and, the opposite, if some parts of their teaching should be changed and more addressed to what is tests (Messick, 1996). These changes may be recommendable in situations when different teachers evaluate different students with the same test. For instance, when the same course is divided into different sections due to a large number of registered students or when two or more teachers share the same group. These changes will help the students to focus in their learning and testing process. For example, if a tourism instructor is using a grammar-centred approach it would not probably be suitable to assess the content of a long composition if demanded by the college, and if so, the teacher may want to reconsider the teaching activity.

Conclusions

Most instructors in ESP are not aware of many issues that can affect their classroom testing (low stakes) procedures. This paper has presented a succinct study of some of these matters. Among the things to consider in their assessment practices are:

- Watch if a test can be biased by previous subject knowledge;
- Seek tasks that really reflect what has been done in class;
- Find the right balance between content and language use;
- Include the traditional skills of listening, writing, reading and speaking activities;
- If testing is done through ICT's check that there are no significant differences between the pen-and paper and the computer tests;
- Consider the students' needs when choosing these items. Test items should be aimed to linking the learners' deficiencies and the skills developed in class this may also be related to the washback effect.

By considering these factors, Tourism ESP instructors will be able to improve, rely more strongly in their own testing, and, most likely, revise and improve their own teaching skills.

References

Alvermann, D. E. & Hynd, C. R. (1989). Effects of prior knowledge activation modes and text structure on non-science majors' comprehension of physics. *Journal of Education Research*, 83, 97-102.

Amengual Pizarro, M. (2006) Análisis de la prueba de inglés de la Universitat de les Illes Balears. *Ibérica*, 11: 29-60.

Bitchener, J. Young, S. and Cameron, D. (2005) The Effect of Different Types of Corrective Feedback on ESL Student Writing. *Journal of Second Language Writing*, 14 (3), 191-205.

Chapelle, C.A & Douglas, D. (2006) *Assessing Language through Computer Technology*. Cambridge: Cambridge University Press.

Davies, A. (2001) Logic of Testing Languages for Specific Purposes. *Language Testing*; 18 (2), 133-47.

Douglas, D. (2000) *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.

Douglas, D., & Smith, J. (1997) Theoretical underpinnings of the Test of Spoken English revision project. TOEFL Monograph Series 9. Princeton, NJ: Educational Testing Service.

García Laborda, J. (2006) “PLEVALEX: A new platform for Oral Testing in Spanish” *Eurocall Review*, 9.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.

Pritchard, R. M. O. & Nasr, A (2004) Improving Reading Performance among Egyptian Engineering Students: Principles and Practice. *English for Specific Purposes*; 23 (4): 425-445.

Salmani-Nodoushan, M. A. (2003) Text Familiarity, Reading Tasks, and ESP Test Performance: A Study on Iranian LEP and Non-LEP University Students. *Reading Matrix: An International Online Journal*; 3 (1)

Skehan, P. (1984). Issues in the testing of English for specific purposes. *Language Testing*, 1, 202-220