

Developing Online Tests for Special Language Needs

Martin Beck

E-Mail: beck@mondiale.de

MONDIALE-Testing GmbH (Switzerland)

Information and communication technology has opened a myriad of opportunities for the ELT professional. The instructor can create a customized training program by combining and adapting multi-media teaching material to fit her needs. When it comes to assessment, most tests, even state-of-the-art online testing engines, do not offer teachers the possibility to provide individualized content.

Modern internet-based test frameworks are powered by high-speed databases containing tens of thousands of test items classified in many different categories. This technology enables dynamic test generation. If user-friendly editing and enabling tools are provided to the instructor, along with the necessary guidance, it becomes easy to add content for special purposes to otherwise standardized tests. These items can also make use of all the categorization and calibration facilities available in the test engine.

In the course of the presentation, we will discuss the features necessary in an online testing framework. We will cover the topics of task creation, selection, calibration and validation.

More than any other form of online testing, language testing is confronted by two diametrically opposed demands. The requirement for validity on the one hand and the need for efficiency on the other.

Validity in online testing covers a number of different topics beginning with the content and scope of the test items, extending to repeatability and comparability of the results and finally including issues of security like preventing the use of unauthorized resources (cheating) and authentication of the test candidate; insuring that the candidate is actually the person he purports to be.

The assessment of online tests achieves its efficiency through objective or algorithmic evaluation and for many subjects this approach works well. Language listening and reading comprehension skills lend themselves well to objective evaluation but creative skills such as speaking or writing can only be adequately assessed subjectively, requiring human intervention and potentially impeding efficiency.

The objective of successful online language testing is to attain an acceptable degree of validity at the same time maintaining a level of efficiency commensurate with the expectations of online testing. There are many different types of tests for different objectives, but the test types most widely offered online: placement and proficiency tests. Placement tests are those taken prior to participating in a language training course to assign the new student to the appropriate group. Proficiency tests either test the achievements of a participant in a training program or assess the language skills of an applicant for admission to an institution, employment or transfer. Higher standards must be applied to proficiency tests. An important aspect of validity in the context of proficiency tests is that the results must be easy to interpret and comparable to assessments from other sources. In short they must be mapped onto an accepted standardized scale such as the Common European Framework for Languages.

Another variable in the validity equation is the granularity or the preciseness of the assessment scale. Although the CEF is a helpful tool for interpreting test results, in many cases it does not discriminate sufficiently well to compare test candidates. When a large number of individuals with similar skill levels



must be differentiated, a more exact scale is necessary than for a small number with a broader range of skills. For instance several thousand international students applying for a limited number of openings at a university need to be rated on a much finer scale than a hundred employees being placed into language training classes. An efficient testing strategy must take this into account.

To reach the appropriate balance between efficiency and validity is no easy task. Above all there is no single correct strategy for all situations.

When asked about language testing, most people will immediately think of high-stakes certificate examinations used for selection purposes (academic admissions, career recruitment, immigration). These tests, sometimes called strategic tests, are by far the most prominent examples of language test but tests designed around teaching and learning, tactical tests, are instrumental in bringing students up to the levels required for success in certificate exams. Syllabus-related tests are often used to align students with the appropriate teaching material or to test the achievement during the course. Equally important, are tests that monitor the overall language proficiency and show the big picture. Syllabus-independent tests are necessary to do this.

Taking the teacher's perspective, there are two basic relationships to tests. Either the teacher works for the test by preparing candidates for certificate exams or the test works for the teacher providing insights and tracking learning progress. In order to make effective use of tests, it is essential to select the right test or testing suite to fit the requirements. Especially in a corporate setting, teachers are often required to document the quality of language training by showing proof of progress. Caught up in the process of defining group levels, assigning students and selecting course materials, teachers can easily forget to record the beginning level of the students appropriately. Without this status quo, it is difficult if not impossible to provide exact information about the progress of each student at the end of a course block. Every test is designed for a specific application. Not every test used for general placement purposes delivers detailed information that can serve as a basis for progress monitoring. What is measured and how the test presents the results play an important role. To track progress adequately, the placement and proficiency tests must be compatible, the scoring scale must be granular enough to show average improvement and the tests must assess overall language ability and not mastery of specific tasks presented in a particular book.

Testing and assessment is an integral part of any training program, especially in foreign language instruction because many participants in language training programs will need to take exams for academic admission, job recruiting or career advancement. For this reason, tests should be integrated into the syllabus. Unfortunately many language institutes, particularly in adult education, do not often use tests. This may well be because of the effort involved in authoring, administering and marking tests.

High-stakes tests used by academic admission boards and immigration authorities are well-known in the language training community. These summative tests, which assess a candidate's proficiency across the whole bandwidth of language levels, are not the sort of test that can be integrated into a training program, unless the objective of the course is to prepare for this exam. What institutes need is a cost-effective test which can be tailored to the content of the course and are nowadays more and more tests for special language needs (e.g. ESP)

It should be a formative test that reinforces the learning goals and provides information on areas where the student needs additional work.

Anyone who has taken on the task of developing a test knows how difficult and time consuming test authoring is. And when the test is finished and handed out, often the photocopied test papers do not meet the expectations after innovative, multi-media-oriented instruction. Luckily, the same technology that spawned blended learning provides the solution for state-of-the-art formative testing.



Any computer-based testing system automatically delivers many of the features needed in regards to administration and marking. Depending on the security requirements of the testing situation, the candidate can take a computer-based test anywhere and any time. If a time limitation is to be imposed, most online testing platforms will provide this option. Finally marking is simplified considerably. Formalized questions that test grammatical knowledge can be realized as multiple-choice items and marked automatically. Human intervention is only necessary where production skills are to be assessed. Once each section of the test has been marked, scores are sent by email.

Automated administration and marking, however, are just a fraction of the total effort involved in language testing. The crucial issue is the content for special language needs. This is where most online testing platforms abandon the institution.

The Mondiale Testing GmbH has been involved in language testing for international enterprises for more than 15 years, during which a vast repository of language testing items has been compiled. Through multi-dimensional classification, items which fit the needs individual situations can be readily identified and compiled to a test. Tests can be constructed to be static or dynamic. In addition, if customized content is necessary, the Mondiale Testing authoring teams will work with institutions to create professional quality items.