A Language Proficiency Test that Works

Peter Shih, Hangzhou Panda Education and Technology Group, China
Qiwen Zhou, Hangzhou Panda Education and Technology Group, China

The Asian Conference on Education 2019
Official Conference Proceedings

Abstract
Either out of inadequate technology or for the sake of convenience, most language proficiency tests tend to oversimplify students' diverse needs and provide one size fits all reports. Take the Oxford Young Learners Placement Test, for example. This test includes two sections, namely listening and language use. The latter combines grammar, vocabulary, and language function. A student gets a final score and a corresponding Common European Framework of Reference (CEFR) level at the end of the test. However, such a generic result cannot provide any insight into individualized learning solutions. Consequently, it is of no use for young English language learners (ELLs) in non-English speaking environments. To address this gap, we believe a language proficiency test that truly works should be solution-based, which provides personalized guidance in the service of learning improvement.

Keywords: personalized, efficient, accurate
Introduction

As one of the fast-developing online language education startups in China, we found that current language proficiency tests do not work in guiding learning. As is known, China now has the most significant number of ELLs in the world, who have to take English language proficiency tests to access higher or privileged education. However, current language proficiency tests oversimplify students' diverse needs and provide one-size-fits-all reports. Take Cambridge Test, for example. Each student will get a final score and a corresponding CEFR level at the end of the test. Similar examples are the APTIS and the TOFEL junior test, which also fail to provide diagnoses in detail. Although test-takers receive scores and corresponding CEFR levels in each skill domain, those judgments and descriptions are still one-size-fits-all. In other words, test-takers with similar skill levels do not know what exactly the differences in their language capacities are, and they do not know what learning solutions to come next either. Current language proficiency tests, as we see, also fails to provide learning solutions.

On the meanwhile, for many years, it is the curriculum designers that are providing learning solutions for test-takers. However, most learning solutions provided by curriculum designers are also one-size-fits-all. Over 90% of students of different language proficiency levels in China are taking the same curricula. We think it is time to change.

The change starts with standards. In the past and now, language proficiency tests are designed based on language standards, such as CEFR, CSE (Chinese Standards of English), test-takers get reports that indicate levels and descriptions without learning solutions. Meanwhile, curriculum designers are busy with providing one-size-fits-all learning solutions based on the standards on which the tests are based. The problem is that the solutions provided by curriculum designers do not always meet the diverse demands of test-takers. The effective learning solutions should meet the diverse demands of test-takers and target learners' strengths and weaknesses. To achieve this goal, it is test-developers' responsibility to bridge the gap and provide individualized learning solutions that target at test-takers' strengths and weaknesses.

Main Body

To design a proficiency test that provides individualized learning solutions, we set up standards including language knowledge and language skills based on CEFR and CSE, and we are also designing a computer-adaptive test based on this standards. Tailored learning solutions that target at test-takers' weaknesses are provided at the end of the test.

The Standards

Based on CEFR and CSE, we select the most frequent words and grammar knowledge in daily use and in language tests such as the Cambridge Tests for Young English Learners and the TOFEL Primary Tests. We also establish standards for reading, listening, writing and speaking, which specifically target at Chinese young ELLs’ demands for academic or communicative purpose.
The Test

First, a solution-based language proficiency test provides comprehensive diagnoses. In contrast, language knowledge is underrepresented in reports of current language proficiency tests. In these reports, vocabulary and grammar knowledge is always absent. We believe a solution-based language proficiency test provides specific and constructive diagnoses all around. There has been already much research proving the disadvantages of not testing grammar separately. As Purpura (2004) says, we have no way of knowing what grammatical difficulties learners might experience and providing feedbacks if we test grammar within the reading, speaking, writing, or listening tests. Although whether testing grammar separately yet remains a controversial issue, as an education start-up that has over 20 million strings of learning data, we found the advantages of testing grammar and vocabulary separately from language skills. Based on several pilot tests on a sample of 156,221 young English learners aged from 4 to 12, we find young ELLs have various language acquisition paths both in language knowledge (vocabulary and grammar) and skills (listening, speaking, reading and writing). For instance, learners’ grammar knowledge is not always in sync with their lexical range. For any individual, each aspect of the test results may fall into different categories, sometimes with an enormous gap. Two cases are shown below:

In case one, students A has similar grammatical problems in speaking and writing. These repeated problems are third-person singular, running-on sentences, and sentence fragments. In case two, student B has low performance in grammar knowledge and writing.
We found out it is always the similar problems that impede student A’s performance in speaking and writing; it is a lack of grammatical knowledge, not vocabulary, that leads to student B's low performance in writing. We have a considerable amount of data that makes us believe testing grammar and vocabulary separately helps us know exact difficulties ELLs may experience in reading, writing, speaking, and listening, and provide detailed feedback. The data also indicates that solutions provided by many test developers might be either too general or unreliable because learners' vocabulary and grammar levels are usually underrepresented or assumed identical to their four skills.

Second, a solution-based language proficiency test provides tailored learning paths. Current language proficiency tests are judgment-oriented. As said before, a learner cannot get any learning insight from a CEFR level or a score, or any general level descriptions. We believe a solution-based language proficiency test is learning-oriented and provides learners with the exact daily or monthly plans, including what they still have to learn. For example, on our platform, our students, after test, are expected to be provided with a personalized learning solution per day based on their weaknesses and interests.

In the meanwhile, we are always trying to find out how a personalized learning path impacts students’ learning behavior and results. From our A/B tests, we found with a personalized learning path, student attendance rate is more stable, while in-class performance (exercise scores and engagement) is better than those with a one-size-fits-all learning path.
Third, a solution-based language proficiency test predicts what to learn next. Current language proficiency tests roughly report what learners are already capable of. As we see from a TOFEL Primary test report, students only get a list of can do. Can do is not the start point for learning. Instead, cannot do is. Therefore, we believe a solution-based proficiency test predicts what to learn next. For example, a test-taker, after the test, is immediately provided a report which indicates the 497 words, 22 grammar, 20 reading skills, and 13 writing skills he or she has to learn next.

Take vocabulary test, for example, based on 156,221 students' performing data on 1300 most frequent words for young ELLs, we analyze and calculate the correlation between words. The result helps us find "Benchmark Words" (BMWs), which can represent a group of words thematically or semantically related. For example, students who know the word elephant always also know words such as cat, dog, pig, to name a few. Our research and findings tell us, benchmark words enable efficient predictions of learners' strengths and weaknesses in lexical resources. The finding of BMWs shortens test time by 91.23% within a maximum deviation rate of 14.78%, which means the accuracy rate is expected to be 85.22% or even higher. Learners' strengths and weaknesses are reported in vocabulary volume, and themes within just 5 minutes.

The Solution

Test-takers are going to receive their overall language proficiency levels based on the standards we set. They are also going to receive detailed reports which indicate what to do next in terms of vocabulary, grammar, reading, listening, speaking, and writing. For example, how many vocabulary words to learn and how many skills they still need to improve in reading, listening, writing, or speaking.

Conclusion

Based on the fact that current language proficiency tests for ELLs are either lack of solution or too time-consuming. We believe each ELL deserves an efficient while accurate language proficiency test. Each ELL also deserves a tailored learning solution rather than an oversimplified skill level. It is test-developers' responsibility to design a solution-based language proficiency test.
Reference


