

Three Challenges in English Language Assessment

by Carol A. Chapelle, Iowa State University
Joan Jamieson, Northern Arizona University



Carol A. Chapelle is Professor of TESOL/Applied Linguistics at Iowa State University. She is the author of *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing, and Research* (Cambridge University Press, 2001). Dr. Chapelle is also the editor of *TESOL Quarterly*.



Joan Jamieson received her Ph.D. from the University of Illinois at Urbana-Champaign and is presently a Professor in the Applied Linguistics Department at Northern Arizona University. Dr. Jamieson is the author of several publications on English as a Second Language assessment and computer-assisted language learning.

Anyone involved in language education today recognizes the importance of testing and assessment. At the same time, however, politicians, parents, learners, and teachers all have different perspectives on why tests are so important and what they can be used for. As a consequence, an important part of English language teachers' expertise includes a more thorough understanding of principles of assessment. However, in articulating those principles, teachers and

researchers alike encounter three thorny issues.¹

1. Defining the purpose of a test

Terms such as "placement," "proficiency," and "achievement" might be used in response to a question about what a test is supposed to be used for. These terms may work for the public, but teachers need a more systematic, thorough, and accurate way of considering the purpose of an English Language assessment because the purpose of the

assessment is critical for choosing or developing a good test. We think of test purpose as consisting of three interrelated concepts:

- One way to look at test purpose is to consider the *inference* that test users want to make on the basis of test scores. For example, we designed *Longman English Assessment* to allow test users to make an inference about examinees' English Language reading comprehension, listening comprehension, and knowledge of written structures. Making an inference about these three aspects of English ability was one way of looking at the test purpose.
- A second way is to consider the *use of the test results*. Typically, a test is given because a decision has to be made about examinees, for example, their readiness to study at a university, their eligibility for a job, or the grade that

they should receive in a course. *Longman English Assessment* is not intended for any of these relatively high-stakes test uses; instead, it is intended to give learners an idea of their level of English ability and to offer level-appropriate advice for improving their English.

- The third aspect of test purpose is any other effect that the test is intended to have, such as serving as a motivating, interesting, and informative experience for learners.

2. Understanding what the test measures

If we say that a test is a measure of reading comprehension, for example, as in the discussion of test inference above, two important questions remain: What do we mean by reading comprehension? What does a specific score on the reading comprehension test indicate about a student's reading comprehension?

The first question has to do with how the test inference is defined. Of course, from a psycholinguist's perspective, many aspects of a reader's linguistic knowledge and processes are called into play during reading comprehension. From the perspective of language assessment, we try to simplify this somewhat to include the idea that linguistic processing of morphosyntax and the lexicon are required as well as comprehension of rhetorical organization and cohesion across different test

types and functions.

The second question has to do with how test performance on the specific test questions relates to what the test taker is actually capable of when reading, for example, an English language course book. Since we were interested in what the test performance would mean relative to English language courses, we developed most of *Longman English Assessment* through the use of materials from English language textbooks. We defined the types of questions we would ask on the basis of a detailed — but not too detailed — description of what it would take to understand the texts at various levels.

3. Checking on validity

Everyone seems to agree that they want *valid* English language tests, but as teachers we need to have a clear idea about what this means. In fact, assessment specialists see the expression "valid test" as misleading, because it implies that validity is an inherent quality of a test. A test is not in and of itself valid or invalid; instead, validity refers to the *use* of a test. For example, a test that we develop to assess ability to use complex sentences in an essay might be valid as part of an achievement test in an English language writing course. In other words, that test use would be valid. However, if the same test were used to place students into a particular level of reading comprehension, that test use would be recognized as invalid. The test was the same in both cases, but the validity depended on its use.

The idea that *validity refers to test use* is critical because it means that English language teachers are often best able to comment on validity relative to their own test programs and classes. However, to do so, it is important that they understand the characteristics of tests that help to increase the validity of their use — characteristics such as reliability and authenticity. In developing *Longman English Assessment*, we were concerned about these test characteristics, but the final analysis of the validity of its use needs to be based on the teacher's own understanding of these characteristics in view of its use.

These three testing challenges, as well as other questions about what and why we teach the way we do, are magnified when instruction and assessment take place through technology. Because of the commitment of time and resources involved in constructing a computer-based assessment, developers take great care to define the test purpose, interpret the score meaning, and consider evidence concerning the validity of the test. As a consequence, the use of carefully designed computer-based materials should help teachers to reflect on these three issues in all of the assessments they use.

Find more papers on technology, English language learning, and English language assessment at www.public.iastate.edu/~carolc.

¹ In describing these issues, we are using the terms "test" and "assessment" interchangeably because the basic principles we are describing pertain to the full range of what people refer to as tests and assessments.