

“Creating Meaningful Classroom-Based L2 Exams”

By: Ann Frances Draemel

Abstract

How does one design an L2 criterion-referenced exam that not only encourages students to learn and apply the material taught in class, but also provides meaningful feedback to students and helps instructors assess whether or not they can move forward in their teaching? Creating original classroom-based L2 assessments, like short quizzes, unit tests, and end-of-semester global exams, is a daunting yet necessary task that many English teachers regularly face. Questions, such as where to start, what task types to include, and how to develop a fair scoring system, abound. This presentation will attempt to alleviate the anxieties associated with the exam development process by, first, clearly defining the criteria necessary for test usefulness and, then, by carefully outlining the intricate steps involved in a test’s creation.

Common Misconceptions about Language Testing:

- Believing that there is one “best” test for any given situation
- Having unreasonable expectations about what language tests can do and should be like
- Placing blind faith in pre-made exams and their scoring systems
- Believing that language test creators have magical procedures for creating exams

Important!

Simply creating a test or using a pre-existing one does NOT automatically make it useful for your class’s needs

Six Criteria for Test Usefulness:

- **Reliability**—Is the test that you created a reliable or consistent indicator of the ability that you want to measure? **Ask yourself:**
 1. To what extent do characteristics of the test setting vary from one administration to another?
 2. To what extent do characteristics of the test rubric vary from one part of the test to another, or on different forms of the test?
 3. To what extent do characteristics of the test input or test tasks vary from one part/task/form of the test to another?
 4. To what extent do the characteristics of the expected response vary from one part/form of the test to another?
- **Construct validity**—Are the interpretations of the test scores meaningful and appropriate? **Ask yourself:**
 1. Is the language ability construct for this test clearly and unambiguously defined?
 2. Is the language ability construct for the test relevant to the purpose of the test?
 3. To what extent does the test task reflect the construct definition?
 4. Will the scores obtained from the test help us to make the desired interpretations about test takers’ language ability?

- **Authenticity**—Does the test measure skills that correspond to relevant, real-life target language use tasks? **Ask yourself:**
 1. To what extent does the description of test tasks include information about the setting, input, and expected response?
 2. To what extent do the characteristics of the test tasks correspond to those of target language use tasks?

- **Interactiveness**—Are the students' language abilities engaged in the test? **Ask yourself:**
 1. Does the processing required in the test tasks involve a very narrow range or a wide range of areas of language knowledge?
 2. What language functions are involved in processing the input and formulating a response?
 3. Are the test tasks likely to evoke an affective response that would make it relatively easy or difficult for the test takers to perform at their best?

- **Impact**—Are there potential consequences of the test on students and teachers? **Ask yourself:**
 1. How relevant, complete, and meaningful is the feedback that is provided to test takers?
 2. How relevant and appropriate are the test scores to the decisions to be made?
 3. Are test takers fully informed about the procedures and criteria that will be used in making decisions?
 4. What is the most and least desirable outcome that may result from using the test in this particular way, and how likely is this to happen?

- **Practicality**—Is the test doable? **Ask yourself:**
 1. What resources are required for the design stage, the operationalization stage, and the administration stage of the test?

(Bachman & Palmer, 1996)

Test Development Process:

- What do I need to test?
 - Conduct a needs analysis
- How am I going to test language abilities?
 - Turn target language use tasks into test tasks
- How am I going to give the test?
 - Consider test administration

Conclusion:

The goal of creating meaningful criterion-referenced assessments is to equip English learners with the necessary linguistic means and methods that will ultimately help them to advance toward successful L2 acquisition.

Speaker Biodata

Ann Frances Draemel holds a Master of Arts degree in TESL / Applied Linguistics from Iowa State University and a Bachelor of Arts degree in English literature and Spanish from Kansas State University. For four years, Ann taught EFL as an ‘auxiliar’ in primary and secondary schools throughout Spain. She is now teaching ESL at Temple University in Philadelphia, Pennsylvania. You may contact Ann at adraemel@gmail.com.

References

- Bachman, L.F., & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Douglas, D. (2010). *Understanding Language Testing*. London: Hodder.
- Fulcher, G. & Davidson, F. (2007). *Language Testing and Assessment: an advanced resource book*. New York: Routledge.
- Jamieson, J. (2011). Assessment of classroom language learning. In E. Hinkle (Ed.) *Handbook of research on second language teaching and learning*. London: Routledge.