A Report on a Piloted Test for English Reading Skills

Meixiu Zhang

Northern Arizona University

Abstract

This article summarizes the development, administration, analytical results, and evaluation of a reading test that was piloted at the Program of Intensive English in Northern Arizona University. Designed as an achievement test, the reading test bases its ten subconstructs on the learning objectives in the course syllabus. Test results were analyzed in terms of item difficulty and item discrimination. Also reliability, the agreement co-efficient, and standard error measurement were analyzed through SPSS 20. The results indicate that the test has an acceptable reliability and slightly low agreement co-efficient. The test has basically fulfilled its intended purpose as an achievement test. The teacher is able to make reliable references about whether students have grasped the reading skills taught in class based on the test results. In addition, the test has exerted beneficial impacts on both test takers and the class teacher.

A Report on a Piloted Test for English Reading Skills

This test project aims to assess the reading abilities of intermediate English as second language (ESL) students enrolled in a Program of Intensive English (PIE) at a public university in North America. The mission of an intensive English program is to equip ESL students with language skills for academic studies in North American universities. Reading skill is by any means one of the most important skills taught in PIE. Therefore, developing a summative test to assess reading abilities is necessary to both ESL learners and the PIE program.

## Research Questions

1. To what extent is the reading test reliable?

2. To what extent is the reading test valid?

## Methods

### Participants

Participants include ten students (three females and seven males) aged from 19 to 28 in a Level 6 class in PIE program at NAU. The participants consist of two Chinese, one Saudi, one South Korean, five Brazilian, and one Taiwanese. All of them study at NAU on a student visa.

### Materials

This achievement test includes three about 600-word passages on academic topics with an average lexile measure of 1280L and an average Flesch Kincaid (FK) of 11.7. Each passage has ten multiple-choice items. Each item will be assigned one point for correct selection, and zero points for incorrect selection.

### Administration

The test was administered in Week 14 Fall 2013 semester in the regular classroom in PIE where students usually had classes. The condition of the classroom was pleasant and free of background noise. Sixty minutes were planned for the test. But all the students finished the test within forty-five minutes.

## Results

The results of all thirty items in this test were analyzed in terms of item difficulty (P) and item discrimination (D) through Excel 2007. The data indicate that four items have a negative D value. Eleven items have zero D values, indicating that both students of higher scores and those of lower scores get the same score in these items. In addition, six items—item 5, item 23, item 14, item 17, item 30, item 12—have a P value of one. Table 1 presents the descriptive statistics of the test. Six items out of thirty were automatically removed from the scale by the SPSS because of their zero variance.

Table 1

*Descriptive statistics for final form items*

| N | K | Min | Max | Mean | *Std. D* | Cronbach's Alpha | *Po* | SEM |
|---|---|-----|-----|------|----------|------------------|------|-----|
| 10 | 24 | 8 | 22 | 15.5 | 3.69 | .63 | .71 | 2.24 |

As shown in Table 2, The Cronbach's Alpha of this test (r=. 63) is slightly lower than the required reliability for a classroom test. Also, the test yields an agreement coefficient of .71.

In terms of consequence related validity, this test has the potential to help improve test takers' academic reading skills due to the detailed and timely feedbacks given to examinees. With regard to criterion related validity, the teacher's positive comments concerning students' test scores verify that the test results generally reflect students' reading abilities in class. Moreover, the explicit description of the construct definition and the operationalization lends

support to the construct validity of this achievement test. Also, students' poor performance on items related to summarizing and making inference certify the researcher's difficulty hypothesis, which is indicative of the construct validity.

## Relevance to PIE

As an achievement test, the reading test based its subconstructs exactly on the learning objects in the course syllabus. The class that participated in this pilot study was supplied with detailed feedback about their performance in each learning objective, as well as their strengths and weakness in academic reading skills. The feedback may lead students to devote more efforts to practicing the reading skills that they inadequately mastered. Scores of this test could be used to make two decisions. First, the teacher could use the scores to make inference about whether students passed or failed in the reading section of this course. Second, the program administrator may use the scores as a partial reference to classify the students in this class into two groups based on whether they have passed the cut score -- those who have mastered adequate reading proficiency for university courses and those who have not.

Related Readings

Alderson, J. C. (2000). *Assessing reading*. Cambridge, UK: Cambridge University Press.

Bachman, L., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests.* Oxford, UK: Oxford University Press.

Davey, B., & Lasasso, C. (1984). The interaction of reader and task factors in the assessment of reading comprehension. *The Journal of Experimental Educational*, *52* (4), 199-206.

Engineer, W. (1977). Proficiency in reading English as a second language. *Unpublished Ph D thesis, University of Edinburgh*.

Grabe,W.(1991). Current developments in second language reading research. *TESOL quarterly*, *25*(3), 375-406.

Levine, A., Ferenz, O., & Reves, T. (2000). EFL academic reading and modern technology: How can we turn our students into independent critical readers. *TESL-EJ*, *4*(4), 1-9.

Stoynoff S., & Chapelle, C.A. (2005). *ESOL tests and testing*. *Alexandria, VA: TESOL*.

Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement*, *25*(1), 47-55.