

Analyzing Vocabulary and Content in a CBI Cumulative Unit Quiz

Veronica Wright and Kevin Gowan

Northern Arizona University

Abstract

This report describes a study conducted in two intermediate level Content-Based instruction (CBI) courses taught at the Program of Intensive English (PIE). The study focuses on analyzing the results of a cumulative quiz based on the two foci of content and vocabulary acquisition. The test had five different types of items meant to reflect typical item types found in exams at university level classes. This was to support one of the overarching goals of the CBI course which was to help prepare students for taking classes at university level institutions. Research focused on determining the validity, reliability of the assessment, and also the correlation between content and vocabulary knowledge in the test results. In order to thoroughly and effectively measure these aspects we created and used a table of specifications. We also used the data to calculate the B-index, item facility, descriptive statistics, frequency charts, and KR-20 formula. The study aimed to find support for the CBI instructional format of teaching as well as the connection between content knowledge and learning and vocabulary knowledge.

Keywords: CBI, formative assessment, item analysis

Analyzing Vocabulary and Content in a CBI Cumulative Unit Quiz

Background

The current definition of Content Based Instruction that the CBI course at PIE follows is relatively new. The idea of using subject matter such as topics or themes based on student interest or need as the main content studied in an ESL course first came around in the 1980s (Snow, 2014). Because of its relative novelty, not all teachers understand the benefits of using such an approach. Snow (2014) argues that content-based instruction that provides opportunities for comprehensible output through collaborative tasks as well as comprehensible input, where the focus is on the meaning of what is being said, fulfills necessary conditions for second language acquisition.

The purpose of our project was to construct, administer, and analyze a criterion-referenced test. This test was to assess the learning of vocabulary and content related materials for two sections of level three content-based instruction (CBI) classes at the Program of Intensive English (PIE). Both CBI sections were taught the same material and therefore were given the same test. We planned to check the validity and reliability of the test and check for a correlation between the scores on the vocabulary and content sections to help support the approach of using content matter to help aide second language acquisition.

Research Questions

In order to have a structure for the project, the authors established the following three questions to guide their analysis of the data collected:

1. Does the test have construct validity?
2. Is there a relationship between scores on the vocabulary section and the content section?
3. Is the test reliable?

Methods

Participants

The participants consisted of two sections of low-intermediate students in level three at the Program of Intensive English (PIE) in Flagstaff, Arizona. Section A consisted of eight students with five female students (one from Kuwait, one from Saudi Arabia, and three from China) and three male students (one from Kuwait, one from Saudi Arabia, and one from China). Section B consisted of seven students with one female student from China and six male students (one from Palestine, one from Saudi Arabia, two from Kuwait, and two from China). Students were all young adults with ages ranging from 18 to the late 20s.

Measures

In designing the test we first created a table of specifications. In creating our table of specifications we listed all the information covered in the unit on medicine and explicitly taught in the class, including all the subtopics and vocabulary. We then listed all the item types included in the test. We then tried to write items that covered all the subtopics of the unit in equal amounts while using a variety of item types in the test. In choosing the item types, we tried to choose item that we felt lent themselves to the types of material and information being tested as well as item formats or questions that they would typically encounter in university level classes and assessment.

There were five parts to the assignment. Part 1 was a fill-in-the-blank vocabulary task consisting of five items. It included a word bank with the five item answers. On reflection we realized we should have put at least one more vocabulary word into the word bank to more effectively test their knowledge of the vocabulary words. Part 2 was a vocabulary task that required the students to use a vocabulary word correctly in a sentence. There were three items.

We asked the students to make sure to use the correct form of the word. Part 3 was a vocabulary task that required the students to match the word to the appropriate definition. There were four items and five possible answers. Part 4 was a multiple choice task concerning content of the unit. There were four items with one key and two distractors each. Part 5 was a short answer section that also focused on the unit's content. There were five items, three of which were two-part items. The items asked students to provide specific detail, expound on specific topics, describe the purpose and function of several medicinal topics, and give a recommendation with supporting details.

Procedures

We used a rubric to grade the assignment. We decided together what to accept and left room for student's individual responses, for example in questions that asked students to give recommendations. We also allowed for grammar and spelling mistakes and focused solely on the comprehension of the unit material. There were a total of 24 possible points; 12 for the vocabulary section and 12 for the content section. We decided that the students should show comprehension mastery so we did not take off points for grammar or spelling mistakes in either the vocabulary or content sections. We met up to discuss the scoring in an attempt to norm.

Results

To calculate the reliability, we used the KR-20 (Kuder-Richardson) formula which takes into account the item facility of all of the items. According to Miller et al. (2013), it is "based on the proportion of persons passing each item and the standard deviation of the total scores" (p. 116). The KR-20 score for this assessment was 0.73. This is a relatively high score and shows that the assignment is, in fact reliable. We also found the KR-20 scores for both the vocabulary and content sections. Keeping in mind that one key way of making an assessment more reliable

is to add more items, it is to be expected that since we are analyzing only 12 items each for the vocabulary and content sections, instead of 24, that the KR-20 scores will be lower for the individual sections. That being said, we saw that the content section is more reliable than the vocabulary section. The vocabulary score is 0.43 compared to the content score of 0.61. We believe this is due to the fact that there were two items in the vocabulary section that we deemed inappropriate for the test whereas there was only one inappropriate item for the content section.

The correlation score when comparing the vocabulary and content section scores (0.60) is fairly strong. With a score like this we can say that there seems to be a relationship between vocabulary knowledge and content knowledge. We cannot say however, the exact nature of the relationship or how or what exactly is causing the relationship but the fact that there is a relationship is a form of support for content-based instruction in the classroom and for setting up a class with a dual focus on content knowledge and vocabulary knowledge.

In order to assess the validity of the construct of the test we first analyzed the test's items in connection to our table of specifications. We found that the test items covered all the different sub topics and vocabulary sections taught in the class evenly with a slight increase in number of test items on the last two subtopics listed, which was done purposefully because these were the only two subtopics that the students had not already been tested on. The item types were also fairly evenly spread out between the different item types for each section, with a slightly higher number of short answer questions for the content section than the multiple choice. We did realize however that our validity could have been stronger if we had used the same type of test items to test students' content and vocabulary knowledge instead of using different types of items to test vocabulary (i.e., fill-in-the-blank, write a sentence, and match the definition) and content (i.e., multiple choice and short answer) information. We also analyzed the empirical data that

described the students' performance on the test. The consistently negatively skewed data clearly indicated the students' ability to learn specific content from the class. This is positive evidence for the use of a CBI instruction model in the classroom.

Relevance to the PIE and Second Language Learning

The purpose of this study was to not only to check the validity and effectiveness of the test we constructed to but also to help find support for the way in which CBI is taught here at PIE. Upon reflection, we felt the assignment we created and administered could be used to accurately assess the students' learning of both vocabulary and content specific to the unit we instructed in our CBI classes. We found that the test items were appropriate for low to intermediate-level ESL learners in testing students' knowledge and were successful in meeting the course objectives of simulating university test settings, though further research would be needed in order to substantiate this claim more firmly. We also found that there are specific things we can do in the future to acquire stronger empirical evidence for our claims such as using the same item types to assess both content and vocabulary knowledge and writing items more carefully in order to assess only content thoroughly covered in class instruction. We must also be more aware of the content we covered in order to have a more accurate assessment of the students' learning. The knowledge and experience we gained will help us to find better support for the correlation between content and vocabulary learning, as well as, provide even stronger evidence for positive benefits of CBI instruction in ESL classes.

References

- Bae, C. (2007). The acquisition of vocabulary in content-based instruction: A case study of 13 EFL adult learners. *Sophia Linguistica*, 55, 71-95.
- Gardner, T. H. (2009). Vocabulary integration in a content-based international business English program. *Latin American Journal of Content & Language Integrated Learning*, 2(1).
- Miller, M. D., Linn, R. L., & Gronlund, N. E. (2013). *Measurement and assessment in teaching*. (11th ed.). New York, NY: Pearson.
- Snow, M. A. (2014). Content-based and immersion models of second/foreign language teaching. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 438 —454). Boston, MA: National Geographic Learning.
- Stoller, F. L. (2002). Promoting the acquisition of knowledge in a content-based course. In J. Crandall & D. Kaufman (Eds.), *Content-based instruction in higher education settings* (pp. 109–123). Alexandria, VA: TESOL.