Task-based Language Assessment in an Intensive English Program

Michael Jordan, Tatiana Shulyateva, and Ozge Yol

Northern Arizona University

Abstract

This research investigated how effective the rubric that the researchers created was for task-based assessment of speaking. It also investigated whether a speaking task the researchers created was effective in terms of promoting interaction. The researchers designed a task similar to a real life situation, and a rubric to measure the performance of learners on this task. The task asked the participants to come to an agreement about opening a business in Flagstaff, Arizona. The results showed that the task was successful in promoting interaction between the participants and they were able to engage in meaningful interaction. However, the rubric the researchers created was not successful and needed to be revised.

Key words: *rubric, speaking task, task-based language assessment*

Task-based Language Assessment in an Intensive English Program

Background

Task-based language teaching allows students to practice a controlled form of real-world communication. The primary focus for performance assessment is to get an accurate picture of students' communicative abilities, and to go beyond learning/testing situations to real-life communication. Jabbarpoor (2011) defined TBLA as consisting of three parts: (a) it's based on tasks, (b) the tasks are authentic, and (c) they should be rated by qualified judges. TBLA originated as a practical way for teachers to test listening, speaking, reading, and writing. Ke (2006) emphasized the importance of integrating a mixture of those skills when performing a TBLA (p. 213).

There are three types of tasks that students can perform that range from easy to difficult (Ke, p. 221). When ranked from easiest to most difficult, they are: (a) static task, (b) dynamic tasks, and (c) abstract tasks. Gyson & Van Avermaet (2005) explained that it is best to use abstract tasks because students then have an opportunity to express themselves as individuals. For assessing speaking, it is useful to see the differences between spoken-like and written-like language as a continuum, with highly oral language at one end and highly literate language at the other. In addition to grammar, oral and literate speech differs in their pronunciation and choice of vocabulary. Test designers can vary things like planning time and the kinds of speaker roles and role relationships that they include in the tasks. In addition, speakers are expected to show that they are able to perform like a natural speaker, and know "phrases and strategies for creating time to speak" (Luomo, 2005) such as fillers and hesitation markers or whole expressions. These

should be the points for consideration while designing an assessment scale for the task based assessment.

Ke (2006) proposed many advantages for using TBLA: (a) it is closely related to the activities the students performed in class while learning something new (p. 215), (b) it gets immediate results about how well the students understand what was just taught, and (c) it can be used as a post-task activity where students would not feel as if they are being tested. However, there are some disadvantages. Mislevy et al (2001) mentioned that one of the issues is that TBLA measures a person's knowledge of how language is used appropriately, even when the task is only to measure the person's knowledge of language. Ke (2006) found that when scoring task-based language assignments, it is best to use analytic scoring rather than holistic scoring because students are given scores based on a variety of components.

Research Questions

Before administering the task, the researchers were looking to answer the following questions:

1.   How effective is the rubric that we created for a paired speaking task?

2.   How effective was this task in promoting interaction?

Methods

The participants were 22 students in the level three Listening and Speaking class. The students were 17 to 28 years old from the Middle East, Asia, and South America.

Students were asked to complete a communicative speaking task in pairs. They were informed that they would open a business in Flagstaff, and they need to reach an agreement on which business they should open. They were given two minutes to choose two business options among the six different options provided (e.g., restaurant, flower shop, or nightclub) and get

prepared for the discussion. They knew that their partners had different options. It was

emphasized on the task sheet that the students needed to do the following:

1.   State what business they need to open

2.   Provide reasons why their business idea is good

3.   Provide reasons why their partner's business idea would not work

4.   Ask to repeat or rephrase if they did not understand their partner

5.   Make sure that their partner understood them.

Each group was given a voice recorder and the speaking prompt. However, they were not

provided with a scoring rubric. Each student was evaluated by three raters.

Results

To answer the first research question, the rubric we created was not very successful. The

category of Interaction presented the biggest issue. Subcategories one and two of the Interaction

category were designed to evaluate if students were taking turns and were sharing time. The

problem with these two subcategories was that almost all students were taking turns and sharing

time. Another problem was that they were measuring very similar constructs and it was hard to

distinguish between them. Subcategories three and four were designed to evaluate how well

students can ask for and provide clarifications. Miscommunication occurred between only one

pair, and it was resolved. Thus, raters had to give perfect scores to almost everyone. Subcategory

one in the Task Completion category also presented a problem because it was too easy and 100%

of the students got full scores. This subcategory was evaluating whether students expressed two

business ideas as they were asked to.

To answer the second research question, we believe the task was successful in promoting

interaction because all students recorded at least 2 minutes of conversation. There were several

reasons why the task was successful. First, students had to negotiate and interact in order to complete the task because they could not do the task on their own. Besides, students had to go through several "stages" before they could complete the "final" task (i.e., come up with one business). Second, the topic students were talking about was familiar to them. They just finished a unit on business and entrepreneurship and were familiar with some stages and procedures of opening a business. Finally, all of the students were currently living in Flagstaff and, thus, could find arguments why one or another business could work or not (as observed from students' responses).

<div align="center">**Relevance to PIE**</div>

Although the results of the study may not be applicable to all the PIE students and ESL learners enrolled in Listening and Speaking classes in the USA, this study may help language teachers in both PIE and the USA to plan their task based assessment accordingly. PIE also employs a similar task based assessment in level 3 students, so both the design and results of this study may help instructors and assessment developers to foresee the potential problems, and overcome these in advance. For example, the results showed that raters had difficulty in reaching a consensus due to the design of the rubric, so the inter-rater reliability was low.

The task itself encouraged students to practice speaking with their pairs, so it was beneficial for the learners. In addition, they had a chance to recycle the vocabulary and the topics that they had learned in class because the task was designed based on the unit covered in the syllabus.

References

Gyson, S., & Van Avermaet, P. (2005). Issues in functional language performance assessment: The case of the certificate Dutch as a foreign language. *Language Assessment Quarterly*, *2*(1), 51—68.

Jabbarpoor, S. (2011). TBLA: Difficulty-based task items sequencing and task-based test performance. *International Conference on Languages, Literature and Linguistics*, *26*, 219—223.

Ke, C. (2006). A model of formative TBLA for Chinese as a foreign language. *Language Assessment Quarterly*, *3*(2), 207—227.

Luoma, S. (2008). *Assessing Speaking*. Cambridge Language Assessment. (Kindle Locations 210-212). Kindle Edition.

Mislevy, R., Steinberg, L., & Almond, R. (2001). Design and analysis in TBLA. *Language Assessment Quarterly*, 1—29.