Synthesis Writing Rubric Development for Content-based Instruction

Nick Velde, Nadege Netta, and James Carpenter

Northern Arizona University

Abstract

23 PIE students enrolled in the Level 4 Content Based Instruction (CBI) class at the PIE completed synthesis essays as the culminating project for the first unit of the semester. This project attempted to develop an alternative scoring rubric for these essays by defining the construct of "synthesis writing ability" and applying it to the CBI project. The project description and CBI course materials were used to develop the criteria on the rubric, and all the essays were scored by the three members of the assessment development team. In the end, the rubric was found to be internally consistent, but the inter-rater reliability was not particularly strong, nor did the sub-categories on the rubric correlate widely enough apart to support the notion that they were testing substantially different aspects of the construct.

Background

Students were aged 18 to 30, and spoke either Chinese or Arabic as their first language. Students were enrolled in two sections of an intermediate level CBI class at Northern Arizona University's Program in Intensive English (PIE). The PIE classification for these students was Level 4. Students were prepared for the synthesis essay throughout the duration of the unit, which lasted for three weeks. All students were familiar with the expectations and nature of the writing task at test time.

The test task was essentially identical to the TLU domain described above—i.e. writing a synthesis essay in an academic context. However, the sub-categories presented on the rubric were not necessarily the only categories possible for assessing the construct of "synthesis writing ability." The scoring rubric was designed according to the table of specifications, and consists of four bands: "textual organization and cohesion," "register knowledge," "explains the procedure of gather, process and report," and "synthesizes information." Also as described in the table of specifications, each band could receive a maximum score of 3 and a minimum score of 1, and thus students could receive a maximum total score of 12 and a minimum total score of 4. The cut score was set at 7 out of 12. Cut scores for each band on the rubric were not set because of the size of the scale.

Methods

Descriptive statistics, item analysis and Cronbach's alpha were calculated to determine the internal consistency of the rubric and inter-rater reliability.

Results

The rubric performed well in some ways, and not particularly well in others. The high inter-rater reliability and low SEM suggest that the rubric was relatively easy to use by raters, as far as scoring the items consistently. However, in some cases, the utility of the rubric made raters too consistent, as was the case with rater 1, who gave the same score to every student on the "register knowledge" band. At the same time, the high correlation between sub-constructs is troubling and suggests that many of the criteria defining the different sub-categories on the rubric were actually targeting the similar things. This is supported by qualitative data obtained through the debrief session conducted by the test development team about the efficacy of the rubric. All three of the raters identified the "register knowledge" and "synthesizes information" bands as assessing very similar elements.

Relevance to PIE

In the same way, it is not possible to say that the test achieved its intended purpose in that inferences drawn from the test results do not necessarily generalize to academic contexts outside the CBI class. Also, the test scores cannot be interpreted as an accurate measure of synthesis writing ability, and as a result the consequences (i.e. the effect on student's grade, motivation, etc.) are not justified or beneficial to the students. There may be some benefit to teachers at the PIE, in the sense that the failings of this rubric may help future CBI teachers and assessment development teams not to repeat the same mistakes, however it cannot be said that the test is a valid or reliable means of assessing synthesis writing ability as articulated by the CBI project description.