Abstract

Background: Criterion-referenced interpretations of tests are highly necessary, which usually involves the difficult task of establishing cut scores. Contrasting with other Item Response Theory (IRT)-based standard setting methods, a non-judgmental approach is proposed in this study, in which Item Characteristic Curve (ICC) transformations lead to the final cut scores. Method: eCat-Listening, a computerized adaptive test for the evaluation of English Listening, was administered to 1,576 participants, and the proposed standard setting method was applied to classify them into the performance standards of the Common European Framework of Reference for Languages (CEFR). Results: The results showed a classification closely related to relevant external measures of the English language domain, according to the CEFR. Conclusions: It is concluded that the proposed method is a practical and valid standard setting alternative for IRT-based tests interpretations.

Keywords

Standard setting, item response theory, computerized adaptive testing, criterion-referenced testing.