Abstract

Assessing English Writing in the Austrian Educational Standards Baseline test. The Case of Coherence, Cohesion and Task Achievement.

Florian Pibal

Guenther Sigott

Alpen-Adria-Universität Klagenfurt

This contribution focuses on the impact of errors on the ratings that performances receive on the analytic rating scale in use since 2009. The focus is on the distinction between the *Coherence & Cohesion* dimension and the *Task Achievement* dimension. In order to address these issues, an error analysis of 100 writing performances was carried out. Errors were identified in terms of linguistic category (grammar, lexis, orthography and punctuation, discourse) and surface-structure alterations (addition, misformation, misordering, omission) (cf. Ellis and Barkhuizen 2005; Sigott et al. 2016). The ratings awarded to each performance by a panel of trained raters were submitted to Many-Facet Rasch Analysis (Eckes 2011). The resulting *Fair Measures* were used as dependent variables in correlational and regression analyses. Overall, significant moderate to high correlations were found between the error categories and the *Total Fair Measure*. Discourse-level errors correlate significantly with both *Coherence* and with *Task achievement*. Most importantly, however, coherence errors are a stronger predictor of the *Task Achievement* rating than are pragmatic errors. This indicates that the raters were interpreting the rating dimensions of *Coherence & Cohesion* and that of *Task Achievement* in ways which were not intended by the test constructors. Possibly, raters interpreted flaws in task achievement as breaches in the coherence between the writing prompt and the performance. The results give rise to the question of whether the distinction between *Coherence & Cohesion* and *Task Achievement* should be conceptually sharpened and maintained, or be abandoned. The implications of these findings for the construct validity of the E8 Standards tests as well as for teaching practice are discussed.

References

Eckes, T. 2011. *Introduction to Many-Facet Rasch Measurement. Analyzing and Evaluating Rater-Mediated Assessments.* Frankfurt/M.: Peter Lang.

Ellis, R. and G.P. Barkhuizen. 2005. *Analysing learner language.* Oxford: Oxford University Press.

Sigott, G., H. Cesnik, and N. Dobric. 2016. "Refining the Scope - Substance error taxonomy: A closer look at Substance." In *Corpora in Applied Linguistics. Current Approaches*, edited by N. Dobric, E.M. Graf and A. Onysko, 79-94. Newcastle upon Tyne: Cambridge Scholars Publishing.