Projektname

Predicting Item Difficulty in a Reading Test. A Construct Identification Study of the Austrian 2009 Baseline Reading Test

Projektleitung
Finanzierung
Kooperation

Veröffentlichungen
zum Projekt

Laufzeit

Mag. Klaus Siller

Intern

Siller, K. (forthcoming). Cognitive processes as predictors of item difficulty in the E8 Baseline Reading Test. In G. Sigott (Ed.), *Language testing in Austria. Taking stock.* Frankfurt/M. et al.: Lang.

Siller, K., & Kulmhofer, A. (forthcoming). The development of the E8 Reading Test. In G. Sigott (Ed.), *Language testing in Austria. Taking stock*. Frankfurt/M. et al.: Lang.

2015-2017

Projektbeschreibung

In 2009 educational standards for English as a Foreign Language were implemented in Austria, providing the legal basis for external tests of the English language competence of Austrian pupils after 8 years of schooling. The present dissertation provides evidence of the construct validity of the Austrian E8 Baseline Reading Test by investigating the relationship between characteristics of 120 multiple-choice reading tasks used in this test and their empirical difficulty.

After an extensive outline of the political and historical context of the development of National Educational Standards in Austria and a detailed description of the development of standard-based assessment tasks in general and the Baseline Reading Test in particular, a review of the relevant literature discusses various approaches to language test validation, outlines how reading comprehension has been defined and modelled, and finally identifies and operationalises 55 potential predictor variables that are likely to have an impact on the difficulty of reading tasks.

Results of a simple linear regression analysis show that variables such as *processing operations, metacognitive activities, comprehension level, content, word frequency in the options (k1words), average sentence length, type-to-ken ratio in the item, type-token ratio in the options and lexical density in the item* have a significant impact on the difficulty of the reading items examined. Furthermore, a tendency concerning the impact on item difficulty was found for features such as *the number of different words in the item, the lexical density in the options, the length of the item, word frequency in the options (off-list words), length of the stem, word frequency in the stem (k1-words) and the number of different words in the stem*. However, after combining these features in a multiple regression model, only *processing operations*, such as decoding, establishing propositional meaning, inferencing and building a mental model, and *metacognitive activities*, such as scanning, careful local reading, skimming and careful global reading, remained significant predictors of item difficulty. In addition, a tendency remained for *word frequency in the stem (k1-words)*.

The dissertation concludes with a discussion on how the results of the analyses can help item writers to write items of different complexity and difficulty for future tests, and how teachers can support their pupils in improving their reading skills, thus scoring better in future national reading tests.

PÄDAGOGISCHE
HOCHSCHULE
SALZBURG