

Improving Second Language Proficiency Assessment

A Differential Item Functioning Study

Avi Allalouf

National Institute for Testing and Evaluation, Jerusalem, Israel

Abstract

This study investigates factors affecting knowledge and acquisition of a second language (SL) by examining differential item functioning (DIF) on SL (Hebrew) test items for two language groups: Arabic and Russian speakers. The results are consistent with the literature on English as a SL with regard to performance in grammar and vocabulary. Many items (42%) functioned differentially, indicating a potential threat to validity. The most problematic item type was Sentence Completion. To reduce the number of DIF items included in operational tests, we suggest changing the balance between item types and performing DIF analysis on piloted items. Further research, using the age of the examinees and the length of time they have lived in Israel as explanatory variables, is currently underway. The findings are pertinent to the existing debate regarding the attributes of a critical period in language acquisition. In addition, a special "non-DIF" test form was constructed on the basis of the study's results for purposes of validation. This test form will be administered to Russian and Arabic speakers during 2004.