

A Comparison of Winsteps and 'Dkqi /O I for Vertical Scaling with the Rasch Model

Mark Pomplun

Riverside Publishing Company

MD Hafidz Omar

Riverside Publishing Company

Michael Custer

Riverside Publishing Company

The present study compares vertical scaling results for the Rasch model from BILOGMG and WINSTEPS. The item and ability parameters for the real and simulated mathematics tests were scaled across five grades, second to sixth. The simulated data were based on real data for a series of mathematics tests for Grades 2 to 6. The results from WINSTEPS and BILOG-MG were compared in terms of differences and correlations between estimated item and ability parameters. Generally, WINSTEPS appeared to capture the individual and mean estimates more accurately, and BILOG-MG captured the standard deviations more accurately. However, because of the many possible variations in vertical scaling studies, the generalizability of these specific findings may be limited. More important, the findings illustrate that choice of software, in addition to data collection and scaling method decisions, influences vertical scaling results.

Key Words: vertical scaling • large-scale assessment • Rasch model • IRT software