STRUCTURAL EQUATION MODELING, *10*(2), 276–288

Do Minority Representative Reading Passages Provide Factorially Invariant Scores for All Students?

Mark Pomplun and Md Hafidz Omar
Riverside Publishing

This study investigated the factorial invariance across students from different ethnic groups of a 7th-grade state reading assessment with a narrative passage selected for its Hispanic representativeness. Confirmatory factor analysis was used to assess the fit of 1-, 2-, and 3-factor models to each of the 5 ethnic groups. After a modified 2-factor model was identified as the best fitting, several levels of constraint were investigated. Invariance across all groups was supported for factor loadings and intercepts. However, invariance of the error variances across the groups was not supported. Although the results support the factorial invariance of the scores based on a minority representative reading passage, further research into group differences for the error variances is needed.

As American schools become more diverse, the cultural differences among students raise important concerns about the measurement of reading comprehension. Research into reading comprehension indicates that cultural differences could affect student interest, vocabulary knowledge, and prior knowledge of passages used to assess reading comprehension. One recommendation for fair testing is to create a testing pool with a variety of passages that reflect the experiences and values of different cultures (Murphy, 1998). However, in a large-scale testing program, only a limited number of passages are exposed to each assessed student. This study examined the validity of scores from a large-scale assessment program where only two passages were used to assess reading comprehension and where one of those passages represented Hispanic culture.

Score comparability ensures that the meaning and interpretation of the test score are the same for all groups of students. For achievement tests, there are three kinds of evidence for score comparability (Willingham et al., 1988). These are reliability, factor structure, and item functioning. Although observed score