
Two techniques for detecting differential item functioning ("DIF") of test items are compared: (1) item response theory (IRT), using the three-parameter model; and (2) Mantel-Haenszel chi square techniques (MHCS). The steps necessary for identifying differentially functioning items are identified for both approaches. Using data from the 1988 Maryland Test of Citizenship Skills (MTCS), the following parameters were investigated: (1) the stability of the MHCS statistic across sample sizes; (2) the stability of the DIF index (MH Alpha) in the MHCS approach across several score groups; (3) the correlation between IRT DIF indices and MH Alpha; and (4) agreement between the IRT and MHCS techniques in identifying biased items. The MTCS was administered to about 50,000 ninth-graders in January and February of 1988. For IRT purposes, random comparison groups of 1,000 students each were created, while the MHCS procedures used samples of 1,000, 750, 500, and 200 examinees. The IRT DIF procedure identified four items in male/female comparisons and three items in white/black comparisons in the MTCS that have significant unequal probabilities of a correct response. While correlation results indicate that the two techniques disagree regarding item bias, agreement in terms of total hits indicates that the MHCS technique is an adequate substitute for the three-parameter IRT approach if the sample size is at least 750. In terms of the percentage of items identified as biased, the techniques appear to be equally good in detecting DIF. Five graphs and 11 tables present study data. A 35-item list of references is provided. (SLD)
Detecting Differential Item Functioning Using IRT and Mantel-Haenszel Techniques: Implementing Procedures and Comparing Results

HEIBATOLLAH BAGHI
STEVEN F. FERRARA

DIVISION OF INSTRUCTION
Maryland State Department of Education
200 West Baltimore Street
Baltimore, Maryland 21201

Paper presented at the Annual Conference of the Eastern Educational Research Association

Clearwater, Florida February 1990