ABSTRACT

Item analysis is a staple of educational and psychological measurement. And within the bailiwick of item analysis the most fundamental criterion is item difficulty. In application several measures of item difficulty exist. For its part, the multiple-choice form of question has endured for nearly a century, with banks of questions accompanying many, if not most, introductory level textbooks. Using the multiple-choice question examination as a context, the present Monte Carlo study compares three common measures of item difficulty and investigates specific examination conditions under which the measures are or are not essentially interchangeable.

KEYWORDS: Item analysis, Multiple-choice questions, Item difficulty, Monte Carlo simulation