

Assessing the Impact of Guessing on Coefficient alpha and Reliability

Introduction

In multiple-choice tests, guessing can influence one of the major psychometric properties of a test such as test score reliability. This study examined the degree of the impact of guessing in multiple-choice tests on reliability and Coefficient alpha. Under the situations on which this study focuses, precisely speaking, Coefficient alpha is no longer a reliability index, thus reliability was estimated directly through the ratio of the true score variance to the observed score variance. Simulations were conducted using the item response theory (IRT) approach. The IRT-based simulation approach permitted relaxing the use of restrictive assumptions such as parallel tests, homogeneous inter-item correlations, homogeneous item difficulties, and homogeneous guessing.

Some major features of the simulation approach employed in this study include: first, modeling of partial guessing behaviors which is more realistic than the random guessing behavior modeling; second, computational efficiency because of no need of generating actual item response data. All simulations and computation of Coefficient alpha and reliability were facilitated by expressing classical test theory (CTT) item statistics, Coefficient alpha, and reliability in terms of an IRT model parameters; third, modeling of nonlinearity which exists in the binary item response data.

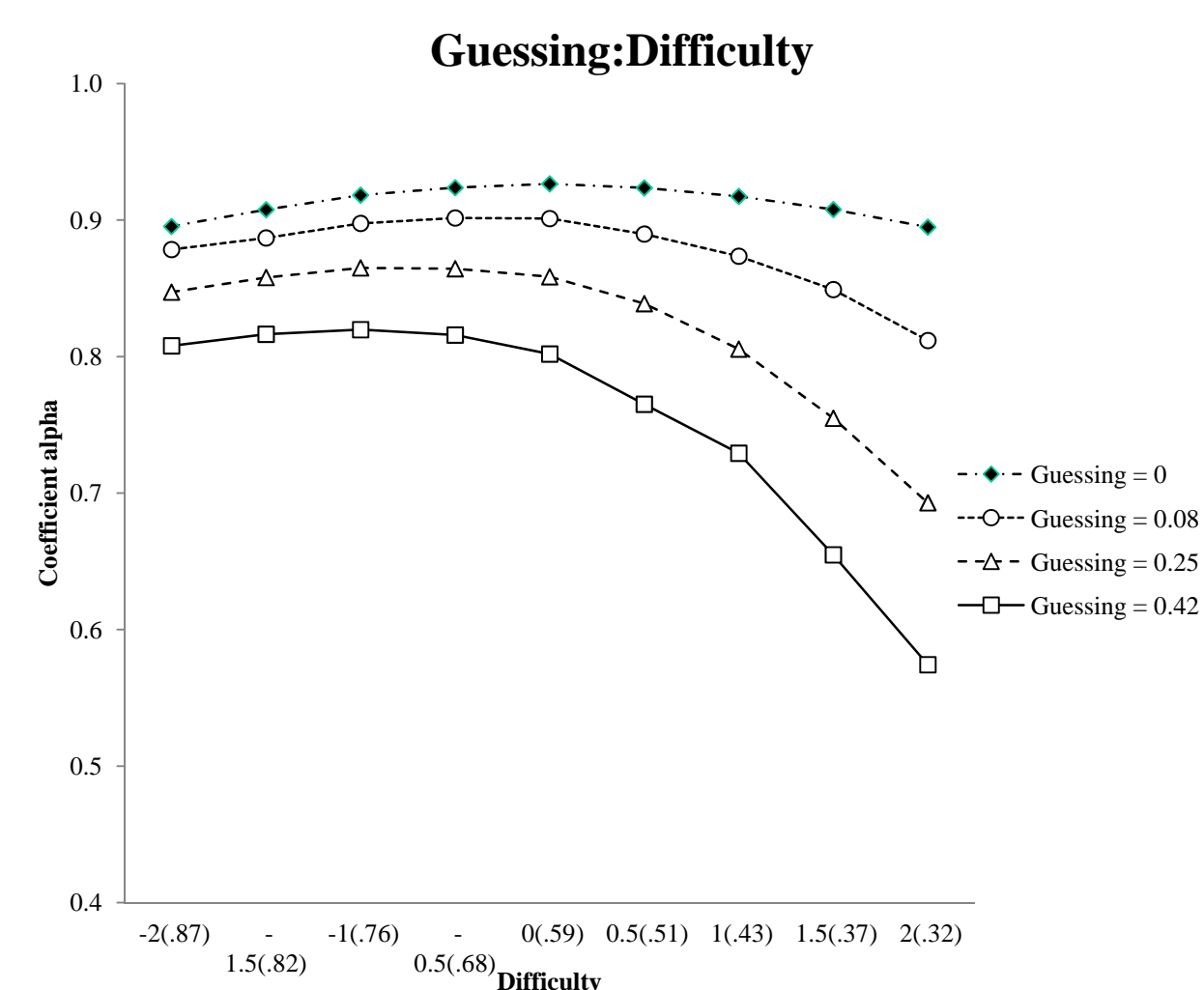
Study Design

The factors considered in the study were test length (small, medium, and large), item difficulty variations (easy, medium, and hard), item discrimination variations (low, medium, and high), and guessing levels (little, medium, and high). The major foci of the study were on the effects of guessing and its interaction with other factors.

Results

1. The differences between reliability and Coefficient alpha were small or very small
2. The general negative impact of guessing on reliability and Coefficient alpha as a main effect was second-largest after the test length effect.
3. Interaction effects between test length and guessing and between guessing and test difficulty were noticeable.

Plot of the Interaction effect between overall test difficulty and guessing levels on Coefficient alpha



Discussion

- Although this study accommodated many important aspects of real test characteristics, limitations exist, for example, the distributions of item difficulties employed in this study, which were assumed to be from a normal distribution.
- One of the future directions may include relaxing these distributional characteristics of item and person ability distributions and employ non-normal sampling distributions (e.g., skewed ability distributions) to examine the impact of guessing under such situations.

Author contact information: Ins Paek

Department of Educational Psychology & Instructional Systems

Email: ipaek@fsu.edu