Abstract

This thesis aims to introduce the use of computerized adaptive testing (CAT) – a novel and ever increasingly used method of a test administration – applied to the field of Kinanthropology. By adapting a test to an individual respondent's latent trait level, computerized adaptive testing offers numerous theoretical and methodological improvements that can significantly advance testing procedures.

In the first part of the thesis, the theoretical and conceptual basis of CAT, as well as a brief overview of its historical origins and basic general principles are presented. The discussion necessarily includes the description of Item Response Theory (IRT) to some extent, since IRT is almost exclusively used as the mathematical model in today's CAT applications. Practical application of CAT is then evaluated using Monte-Carlo simulations involving adaptive administration of the Physical Self-Description Questionnaire (PSDQ) (Marsh, Richards, Johnson, Roche, & Tremayne, 1994) – an instrument widely used to assess physical self-concept in the field of sport and exercise psychology. The Monte Carlo simulation of the PSDQ adaptive administration utilized a real item pool (N = 70) calibrated with a Graded Response Model (GRM, see Samejima, 1969, 1997). The responses to test items were generated based on item parameters and pre-specified true latent values of physical self-concept. The Monte Carlo simulation was designed to compare the number of administered items from PSDQ (test length) and accuracy of estimated latent levels of physical self-concept. The simulations also allowed for a comparison of different latent trait estimation methods, item selection procedures and other frequently used settings within current CAT research.

Results have shown that CAT can successfully be applied as a method of reducing test length when measuring physical self-concept using the PSDQ items. In particular, adaptive administration saved on average between 50% and 90% of the PSDQ items (depending on required measurement precision), while the obtained latent trait estimates were relatively unbiased and very similar to those based on the full version of the PSDQ. The maximum likelihood latent trait estimation, expected a posteriori estimation with uniform prior, and expected a posteriori estimation with standard normal prior distribution were similarly efficient with regard to the average number of administered items in PSDQ CAT. Similarly, comparison of Kullback-Leibler divergence-based and Fisher information-based item selection methods respectively, revealed almost identical results with regard to both the test length and bias of the latent trait estimates.

A possible limitation of this study is that the present findings are based exclusively on Monte-Carlo simulations. Real world applications of PSDQ CAT may however produce slight model and parameter deviations that do not completely conform to simulation findings. Notwithstanding, the promising findings from the simulation suggest a next step would entail evaluating the utility and precision of the PSDQ CAT administration in real testing conditions.

Key words: computerized adaptive testing, item response theory, self-concept, test administration