LARGE-SCALE MEASUREMENT OF 10-11 YEAR OLD STUDENTS' WORD READING SKILLS

Andrea Magyar *, Gyöngyvér Molnár **

* Doctoral School of Education, University of Szeged
** Institute of Education, University of Szeged

Keywords: technology-based assessment; adaptive testing

An appropriate reading level is an essential prerequisite for all further learning (*Csapó* & *Csépe*, 2012), and it is important to monitor its improvement and indicate who has been left behind. *József Nagy* (2006), recognizing the crucial role of the vocabulary and word-reading skills in reading skills development, created a criterion-referenced paper-based test system, which is suitable for the measurement of the knowledge of the most frequent 5,000 colloquial words.

With the spread of computers, the conditions for exploiting the opportunities of computer-based tests as widely as possible have been offered (Molnár, 2010), such as a motivating environment (*Thompson* & *Way*, 2007), the possibility of instant feedback and a more accurate method for the estimation of the persons' ability levels through adaptive testing technology (*Eggen*, 2007; *Magyar*, 2014; *Thompson* & *Weiss*, 2011), which was the aim of this study. During the construction of the adaptive test system the main objective was to retain the original structure of the test, so the measurement of the word-reading skill has been established according to four different perspectives (headwords, inflected words, synonyms and word-meaning reading). From the many possible adaptive structures the four-stage adaptive model has been chosen with five different difficulty levels in each stage. The sample for the study was drawn from Hungarian primary school children at the age of 10 and 11 (N=3,220). The test was constructed and delivered by the eDia online platform by using computer facilities available at schools. According to the results, the system proved to be reliable. The person-separation reliability was adequate both at the level of the entire system (.92) and at the level of the individual dimensions (.73, .85, .91, and .90). The difficulty parameters of the items have been located along a wide range (152 to 804), so the system proved to be suitable for the detection of differences between the grades, and also allowed the diagnosis of the differences that appeared within the grades. For the lower ability level students typically easier, while for the higher level ability students more difficult clusters were administered during the testing process, thus the amount of information obtained during the test gradually improved as each student got the appropriate difficulty level items in the latter phases.

The study has proved that a paper-based word reading test system could be developed into a multi-stage adaptive testing system, and the test system can be used effectively and reliably for the criterion referenced diagnostic measurement of the reading skills of 4-5 graders. The main conclusion of the research is that the test system is suitable to measure the number of words the student is able to read, and also can determine which words are not recognized, even if those words were not included in his/her test version.

G1