



Assessing construct of DIBELS: Measurement invariance of DIBELS subscales, from kindergarten to the first grade



Cody Ding^{a,*}, Ying Liu^b

^a University of Missouri – St. Louis, United States

^b Rowan University, School of Osteopathic Medicine, United States

ARTICLE INFO

Article history:

Received 21 February 2013
Received in revised form 7 November 2013
Accepted 25 November 2013
Available online 24 December 2013

Keywords:

Reading skills
DIBELS
Measurement invariance
Common factor
Structural equation modeling

ABSTRACT

A challenge using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in studying reading growth is that reading skills children exhibit change by age. In order to study growth using changing subscales, it is necessary to examine measurement invariance and measurement structure underlying the different subscales. The purpose of this paper is to examine the measurement structure of the DIBELS subscales, particular measurement invariance. The results indicate that the DIBELS subscales do not seem to have metric invariance but they do share a common factor over time, suggesting that the same construct of reading skills were measured but they manifested in the different fashion over time.

© 2013 Elsevier Ltd. All rights reserved.

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is one of the battery of tests that has been used by many schools to identify children who may be at risk of reading failure, to assess their reading progress, and to guide instruction (Samuels, 2007). Conceptually, it measures students' early literacy skills dynamically with a set of brief probes. Different from curriculum-based measurement (CBM), DIBELS consists of standardized items not based on specific classroom contents and monitors students' progress in an ongoing manner. Tests are administered three times a year, and the type and difficulty of the skills assessed remain the same for a given skill. Typically, children start with low scores and keep improving as they learn. Scores can be compared to benchmark goals that have been empirically validated, thus helping educators identify struggling students and provide appropriate intervention (Good & Kaminski, 2002a, 2002b). Used primarily from kindergarten through third grade with additional measures available for grades four to six, DIBELS includes a developmental sequence of subtests that evaluate phonemic awareness (Initial Sound Fluency or ISF, Phoneme Segmentation Fluency or PSF), alphabetic principle (Letter Naming Fluency or LNF, Nonsense Word Fluency or NWF), accuracy and fluency reading connected text (Oral Reading Fluency or ORF), comprehension (Retell Fluency or RTF), and vocabulary/oral language (Word Use Fluency or WUF) (Good & Kaminski, 2002a, 2002b). This

developmental sequence is based on theoretical and conceptual models of reading fluency positing that fluent word reading is the result of fluency with sublexical processes (Good, Simmons, & Kame'enui, 2001; LaBerge & Samuels, 1974; Logan, 1997).

Operationally, DIBELS is used as a measurement tool for assessing reading skills and growth by more than 15,000 schools nationwide. Many researchers have investigated its effectiveness. For instance, Al Otaiba et al. (2008) used phonological awareness (such as PSF and ISF) and decoding fluency (such as LNF and NWF) to study reading growth during kindergarten. The variate of letter knowledge and phonological awareness was created based on factor analysis to study the relationship between instructions and growth in these skills. The similar study on reading growth was also conducted by Ritchey (2011) and Francis et al. (2008). Several other studies have used DIBELS subtests to identify at-risk students (e.g., Logan & Petscher, 2010) or to conduct formative assessment and evaluation (Kaminski, Cummings, Powell-Smith, & Good, 2008). There are also studies that focus on evaluation of validity of a particular DIBELS subtest in assessing and predicting reading comprehension (e.g., Fuchs, Fuchs, & Compton, 2004; McKenna & Good, 2003; Petscher & Kim, 2011; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008). DIBELS scores have been compared with other standardized test score, particularly statewide assessments (Barger, 2003; Buck & Torgesen, 2003; Wilson, 2005).

Despite of its wide use, there has been a considerable debate with respect to the utility of the DIBELS in predicting reading comprehension (Manzo, 2005). For example, Pearson (2006) stated that DIBELS is the worst thing for the teaching of reading, and DIBELS is not an adequate indicator of reading comprehension

* Corresponding author at: 404 Marillac Hall, University of Missouri – St. Louis, St. Louis, MO 63121, United States. Tel.: +1 314 516 6562; fax: +1 314 516 5784.
E-mail address: dinghc@umsl.edu (C. Ding).

(Pressley, Hilden, & Shankland, 2005). Riedel (2007) found that among the different subtests of DIBELS, ORF was a better predictor of reading comprehension and it seemed unnecessary to administer other subtests. DIBELS subtests such as NWF or PSF are not prerequisite for reading comprehension and an over emphasis on such reading skills may hinder the instruction of overall reading ability (Goodman, 2006). Furthermore, Samuels (2007) argued that the DIBELS tests are not valid test of the construct of fluency or comprehension but rather the construct of speed or accuracy. Although RTF is designed to assess comprehension of the passage, its score reliability and validity were also questioned (Pressley et al., 2005).

Given the current debates about DIBELS as a measurement tool of reading skills, we specifically examined the measurement issues related to DIBELS in this study. Particularly, given that the conceptual construction of reading skills in DIBELS varied across time and not all subscales were used at each occasion, we tested the measurement invariance of these subtests; that is, did these subtests in DIBELS manifest measurement invariance and to what degree (e.g., configural or strong metric invariance)? If not, whether these subtests share the same latent variable (i.e., a common factor) over time? One reason for such study was that this information could help us to determine whether the subscales used at each time formed a latent factor and whether these latent factors with different subscales share a common factor so that these subscales can be used to study developmental trajectory of reading growth over time. In this study, we used the first-order structural model as a way for studying measurement invariance, with the first-order factors representing the construct of the specific reading skill assessed at a particular time. Then, the second-order model was used for testing a common factor model underlying the different subtests.

In testing measurement invariance, researchers have suggested various forms of measurement invariance (Cheung & Rensvold, 2002; Little, 1997; Meredith & Horn, 2001; Vandenberg & Lance, 2000), which included configural invariance, metric or partial metric invariance, residual invariance, strong or scalar invariance, factor variances invariance, factor covariances invariance, and factor means invariance. Sometimes, these measurement invariance forms were also referred to as configural invariance and metric invariance with three different types: weak, strong, and strict metric invariance (Wu, Li, & Zumbo, 2007). In this study, we are particularly interested in testing configural, metric (or partial metric) invariance, and strong invariance. We contend that the utility of construct of reading skills assessed at each time is related to the extent to which these types of measurement invariance are present. It has been suggested that configural invariance is prerequisite to other types of metric invariance (Cheung & Rensvold, 2002), since it may suggest that the construct of reading skills assessed at each time can be considered similar across time. Metric invariance is necessary for the construct of different DIBELS subscales at each time to be manifested in the same way across time. Strong invariance is necessary for the comparison of means of these factors across time, indicating the measurement scales have the same operational definition across time. On the other hand, the invariance with respect to residuals, factor variance, factor

covariance, and factor means may not be essential because heterogeneous variances or covariances are likely to occur in developmental research (Rogosa & Willett, 1985). Moreover, error variances of the same subscale are allowed to be correlated, which indicates possible method variance inherited in the same subscale over time, as suggested by Kline (2010) and Meredith and Horn (2001). Thus, we only tested configural, metric, and strong metric measurement invariance.

Methods

Participants

The participants of the present study were 1233 kindergarteners (53% boys and 47% girls) from different schools in a mid-west state in the U.S. Thirty-seven percent (37%) of these schools were from the urban area and 63% were from the rural area. The participants were predominately White, including 76% White, 19% African American, and 5% multi-racial children. Among them, 13% were identified as being eligible for special education and 69% with free or reduce-priced lunch (i.e., poverty) status. They were from a large evaluation study that investigated the reading development of children during early childhood.

These same children were followed over time from kindergarten (2004–05) through the middle of the first grade (2005–06). The data were provided by the state education agency.

Measures

The primary measure used to assess reading progress was the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002a). For this cohort of young children, there were a total of five occasions of measurement available so far: three times during kindergarten and two times during the first grade.

DIBELS was designed to assess three key early literacy areas: phonological awareness, alphabetic principles, and fluency with connected text. The measures included the following subscales: initial sounds fluency (ISF), letter naming fluency (LNF), phoneme segmentation fluency (PSF), nonsense word fluency (NWF), oral reading fluency (ORF), retell fluency (RTF), and word use fluency (WUF). Depending on children's grade level, different subscales were used to assess reading progress. According to the publisher (Good & Kaminski, 2002b; Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002), each of the seven scales is a standardized, individually administered test. The difficulty level of the test is comparable to student grade level. The score is based on the number of correct responses in 1 min. Thus, for a given measure, the scores are calibrated on a common scale but for two different scales (e.g. ISF and LNF), the scores are not necessarily on the same scale. Since the official website of DIBELS provides the detailed descriptions of each subtest, we did not repeat them here and readers who are interested in DIBELS can go to <https://dibels.uor-eigon.edu/measures.php> for more information.

The subscales of DIBELS were administered to the participants according to the schedule listed in Table 1. Table 2 shows the means and standard deviations of these subscales across five time

Table 1
DIBELS assessment schedule by grade and time of year.

Grade	Initial sound fluency	Letter naming fluency	Phoneme segmentation	Nonsense word fluency	Oral reading fluency	Retell fluency	Word use fluency
K-Beg	X	X					X
K-Mid	X	X	X	X			X
K-End		X	X	X			X
1-Beg		X	X	X			X
1-Mid			X	X	X	X	X

Download English Version:

<https://daneshyari.com/en/article/372654>

Download Persian Version:

<https://daneshyari.com/article/372654>

[Daneshyari.com](https://daneshyari.com)