Technical Report # 1803

In-Brief: Reliability of the Slope of the easyCBM® Reading Measures

Joseph F. T. Nese

Daniel Anderson

P. Shawn Irvin

Julie Alonzo

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

Note: Funds for this dataset were provided by the U.S. Department of Education: Office of Special Education Programs (OSEP) #H327S150007.

Abstract

This in-brief technical report documents the results from two different analytic approaches for examining the reliability of the slope for easyCBM® reading measures in Grades K-8. Results varied by grade, assessment measure, and the analytic approach. Results patterns are discussed.

In-Brief: Reliability of the Slope of the easyCBM® Reading Measures

Background

The National Center on Intensive Interventions (NCII; https://intensiveintervention.org) evaluates both screening and progress monitoring tools for their technical adequacy, with the goal of helping educators select appropriate tools to meet their needs. Many of the analyses required by NCII are already part of our standard practice in developing assessments and are thus described in detail in the many technical reports we publish as part of our assessment development process. Other analyses, such as the reliability of the slope, reported here, fall outside the scope of our standard technical reports. In this brief, we present the results of our analyses of the reliability of the slope conducted with a population of students in need of intensive intervention.

Methods

Sample

The analytic sample consisted of students who took easyCBM® reading progress monitoring measures during the 2014-2015, 2015-2016, and 2016-2017 school years. All students in the sample were identified by their districts as needing intensive intervention in the specific skill area targeted by the assessments for which their data were included in this study. Data from this study are a subset of a much larger extant data set. The larger data set includes scores for all students in all districts with easyCBM® accounts covering fall of 2014 to spring of 2017. From this larger data set, we included only those students identified as needing intensive intervention who had a minimum of 10 assessment scores for a given assessment measure with a minimum of 20 weeks between the first and last administration occasion. Thus, sample sizes varied by grade and assessment measure administered.

Analyses

We analyzed the reliability of the slope using two approaches, *Pearson split-test correlation analysis* and *reliability of the slope*.

Pearson split-test correlation analysis. For each student, assessments were divided into two data subsets comprised of odd and even numbered tests, respectively, depending on the chronological order in which they were taken. An OLS slope of improvement (growth) was estimated for each data subset and for each student. Table 1, below, summarizes Pearson correlation coefficients as a measure of the strength of association between even and odd numbered test slopes.

Reliability of slope. Reliability of the slope is defined here as the ratio of the true score variance to the total variance. The true score variance is the random slope variance in a mixed-effects growth model (lme4 package; Bates, Maechler, Bolker, & Walker, 2015) in the R software environment (R Core Team, 2018). The total variance is the estimation of total variance of each student's individual slope of improvement (R Core Team, 2018). Table 2 summarizes correlation coefficients as a measure of the strength of association between the true score variance and the total variance.

Results

Results varied by measure and grade, as well as by analytic approach. Table 1 presents the results of the *Pearson split-test correlation* analyses and Table 2 presents the results of the *reliability of the slope* analyses.

Table 1

Pearson Split-test Correlation Analyses Results

Measure	Grade	*n*	Correlation coefficient (*r*)	95% Confidence Interval	
				Lower	Upper
Letter Names (LN)	K	253	.83**	.79	.87
	1	44	.71**	.53	.83
Phoneme Segmenting (SEG)	K	42	.79**	.64	.88
	1	91	.72**	.61	.81
Letter Sounds (LS)	K	237	.84**	.79	.87
	1	199	.60**	.50	.68
Word Reading Fluency (WRF)	K	10	.81	.38	.96
	1	348	.88**	.85	.90
	2	155	.42	.28	.54
	3	60	.42	.19	.61
Passage Reading Fluency (PRF)	1	75	.85**	.77	.90
	2	412	.33	.24	.41
	3	387	.42	.33	.49
	4	163	.26	.11	.40
	5	263	.25	.13	.36
	6	64	.40	.18	.59
	7	11	-.13	-.68	.51
	8	137	.29	.13	.43
Vocabulary (VOC)	2	41	.40	.11	.63
	3	7	.00	-.76	.75
	4	22	.21	-.23	.58
	5	17	.66	.36	.86
	6	7	.56	-.33	.92
	7	1	NA	NA	NA
	8	32	.02	-.33	.36

Note. **Lower bound of the confidence interval around the median correlation estimate falls below 0.50 but meets or exceeds 0.40.

Table 1 cont.
Pearson Split-test Correlation Analyses Results

Measure	Grade	*n*	Correlation coefficient (*r*)	95% Confidence Interval	
				Lower	Upper
Multiple Choice Reading Comprehension (MCRC)	2	138	.34	.18	.48
	3	117	-.12	-.29	.07
	4	42	-.05	-.35	.26
	5	116	.17	-.02	.34
	6	NA	NA	NA	NA
	7	16	.09	-.43	.56
	8	238	.14	.01	.26
CCSS Reading (CCSS)	3	9	.34	-.42	.82
	4	4	-.46	-.99	.90
	5	16	.26	-.27	.67
	6	6	-.54	-.94	.49
	7	4	.25	-.94	.98
	8	22	-.14	-.53	.30

Note. NA = Not analyzed. [**]Lower bound of the confidence interval around the median correlation estimate falls below 0.50 but meets or exceeds 0.40.

Table 2
Reliability of Slope Analyses Results

Measure	Grade	*n*	Correlation coefficient (*r*)	Confidence Interval	
				Lower	Upper
Letter Names (LN)	K	253	.87**	.74	1.00
	1	44	.79**	.50	1.00
Phoneme Segmenting (SEG)	K	42	.83**	.53	1.00
	1	91	.72**	.51	.98
Letter Sounds (LS)	K	237	.86**	.72	1.00
	1	199	.72**	.57	.89
Word Reading Fluency (WRF)	K	10	1.00**	.45	1.00
	1	348	.92**	.81	1.00
	2	155	.75**	.58	.94
	3	60	.54	.29	.87
Passage Reading Fluency (PRF)	1	75	.88**	.65	1.00
	2	412	.62**	.52	.74
	3	387	.56**	.45	.68
	4	163	.53	.37	.72
	5	263	.42	.29	.56
	6	64	.48	.23	.81
	7	11	.03	NA	NA
	8	137	.40	.22	.62
Vocabulary (VOC)	2	41	.07	.00	.42
	3	7	.47	.00	1.00
	4	22	.48	.08	1.00
	5	17	.77	.33	1.00
	6	7	.44	.00	1.00
	7	1	NA	NA	NA
	8	32	.08	.00	.55

Note. NA = Not analyzed. **Lower bound of the confidence interval around the median correlation estimate falls below 0.50 but meets or exceeds 0.40.

Table 2 cont.
Reliability of Slope Analyses Results

Measure	Grade	*n*	Correlation coefficient (*r*)	Confidence Interval	
				Lower	Upper
Multiple Choice Reading Comprehension (MCRC)	2	138	.37	.21	.57
	3	117	.20	.02	.41
	4	42	.28	.00	.72
	5	116	.44	.26	.65
	6	NA	NA	NA	NA
	7	16	.16	.00	.77
	8	238	.27	.12	.43
CCSS Reading (CCSS)	3	9	.72	.11	1.00
	4	4	.08	.00	1.00
	5	16	.27	.00	1.00
	6	6	.23	.00	1.00
	7	4	.67	NA	NA
	8	22	.06	.00	.58

Note. NA = Not analyzed. **Lower bound of the confidence interval around the median correlation estimate falls below 0.50 but meets or exceeds 0.40.

Discussion

Results varied by measure, grade, and analytic approach, with clear patterns emerging. We document reasonable reliability of the slope, whereby the lower bound of the 95% confidence interval around the median correlation estimate falls below 0.50 but meets or exceeds 0.40, for the early/emergent literacy measures in Grades K-1 using both analytic approaches (LN, SEG, and LS) and additionally for WRF and PRF using the *reliability of slope* approach in Grades K-2 and Grades 1-3, respectively. Results for later elementary grades (i.e., generally, Grade 4 and up) and three measures (i.e., VOC, MCRC, and CCSS) were relatively less encouraging. Spanish language measures for all available grades were not analyzable due to insufficient sample size. Given extant data were used for the analyses reported here, we are currently planning a series of studies to better control for sample population characteristics and size and anticipate improved reliability of the slope results for these later grades and measures.

References

Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., & Chang, W. (2018). rmarkdown: Dynamic Documents for R (R package version 1.9). https://CRAN.R-project.org/package=rmarkdown

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01

Grolemund, G., Wickham, H. (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, *40*(3), 1-25. http://www.jstatsoft.org/v40/i03/

Hester, J. & Wickham, H. (2018). fs: Cross-Platform File System Operations Based on 'libuv' (R package version 1.2.6). https://CRAN.R-project.org/package=fs

Müller, K. (2017). here: A Simpler Way to Find Your Files (R package version 0.1). https://CRAN.R-project.org/package=here

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. (2018) psych: Procedures for Personality and Psychological Research (R package version 1.8.4). Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych

Rinker, T. W. (2018). numform: A publication style number and plot formatter (R package version 0.5.0). http://github.com/trinker/numform

Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse' (R package version 1.2.1). https://CRAN.R-project.org/package=tidyverse

Xie, Y. (2018). knitr: A General-Purpose Package for Dynamic Report Generation in R ((R package version 1.20). https://cran.r-project.org/web/packages/knitr/index.html