



DYNAMIC INDICATORS OF BASIC EARLY LITERACY SKILLS 8TH EDITION

DIBELS® 8th Nonsense Word Fluency: Comparison of Administration Directions

Technical Report #26-004

CONTRIBUTING AUTHORS

Havisha Khurana, PhD candidate,
Quantitative Research Methods in Education (QRME),
University of Oregon

Gina Biancarosa, Ed. D.
Associate Professor and Ann Swindells Chair in Education,
Department of Special Education and Clinical Sciences,
College of Education,
University of Oregon

Janet Otterstedt, MS.,
Research Assistant and Project Coordinator, Center on Teaching &
Learning, College of Education,
University of Oregon

Patrick C. Kennedy, Ph.D.,
Research Associate and Director of Data Management and Analysis
Group, Center on Teaching & Learning, College of Education,
University of Oregon

TABLE OF CONTENTS

Abstract	4
DIBELS® 8th Nonsense Word Fluency: Comparison of Administration Directions	5
Method	10
Participants.....	10
Measures.....	13
Dependent Variables.....	13
Independent Variables.....	13
Covariates	14
Procedures.....	14
Analyses.....	14
Results	15
Discussion.....	21
References.....	22

ABSTRACT

The Nonsense Word Fluency (NWF) subtest of the DIBELS® 8th assessment battery measures a student's ability to decode and read phonetically regular letter combinations that are not actual English words resulting into two scores - Correct Letter Sound (NWF-CLS) and Words Read Correctly (NWF-WRC). Given the weighted contribution of the scores to the composite score and the timed nature of the NWF task, we want to avoid signaling to students that they should both sound out each letter and read the whole word because a correct response using either approach results in same score but, potentially, requires a different amount of time. In response to user feedback on the operational NWF administration directions, the current study explores the effects of two versions of administration directions on NWF-CLS and NWF-WRC scores using a within-student randomized control trial (RCT). Employing a series of two-level random-intercept linear regression, we find no average effect of the revised directions on NWF scores although the revised directions have higher face validity. Implications and limitations of the study are discussed.

DIBELS® 8th NONSENSE WORD FLUENCY: COMPARISON OF ADMINISTRATION DIRECTIONS

DIBELS® 8th Edition is a commonly used assessment battery of early literacy skills in kindergarten through Grade 8 in the United States. One of the subtests, Nonsense Word Fluency (NWF), is a standardized, individually administered subtest given to students in kindergarten through Grade 3 to measure student's ability to decode and read phonetically regular letter combinations that are not actual English words. For example, "tup" is phonetically regular and therefore decodable but is not a real word. Two scores, Correct Letter Sounds (CLS) and Words Read Correctly (WRC), are recorded during the NWF administration, and both contribute to the DIBELS® 8 composite score.

The CLS score is more heavily weighted than the WRC score in the composite formula. In addition, NWF is a timed task lasting only 60 seconds. Some users have hypothesized that the current directions give students the impression that they should both sound out individual letters and then read the whole word. For example, a nonsense word like tup could be sounded out and then blended (i.e., "/t/ /u/ /p/ tup") or simply blended (i.e., "tup"). For a given word, a correct response using either approach results in the same number of points for both scores, but the former takes more time than the latter, reducing the total number of words a student can attempt in 60 seconds. Thus, the behavior that maximizes scores on NWF is to blend the nonsense words without sounding them out, assuming a student has the skills to do so.

Therefore, the current study explores the effects of two versions of the directions on NWF-CLS and NWF-WRC scores. Specifically, we tested the existing directions and a revised version of the directions designed to make the optimal behavior clearer. The key difference between the versions (see Table 1 for details) is that the revised directions model the blending approach before sounding out as acceptable behaviors and clarify that the sounding out approach can be used when needed.

We employed a within-student randomized control trial (RCT) in which each participant completed two NWF forms: one with the existing directions and one with the revised directions. Form order and directions were counter-balanced and assigned randomly for each student. We investigated the following research question: Does the performance on both NWF scores differ depending on the version of administration directions used, after accounting for repeated observations within students, grade-level, school, and form order? We hypothesize that there is no difference in NWF scores between the original and revised administration directions, on average. If there is a difference in average NWF scores by administration directions, it would suggest a systematic difference in assessment scores, reflecting that the directions influence the construct being measured.

Table 1.
Original and Revised Directions for NWF

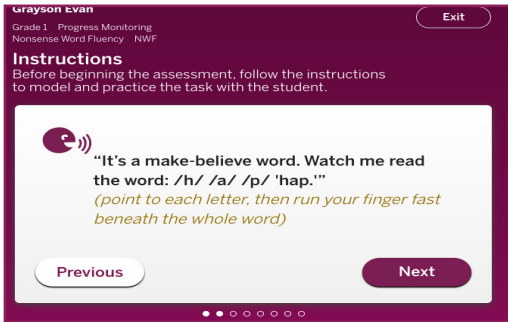
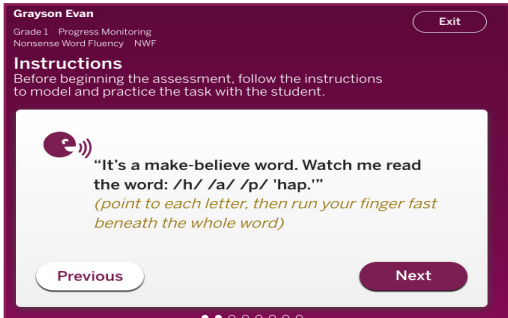
#	Original Directions	Revised Directions
1		No change
2		“It’s a make-believe word. Watch me read the word: ‘hap.’” (run your finger fast beneath the whole word)

Table 1. *Cont'd*

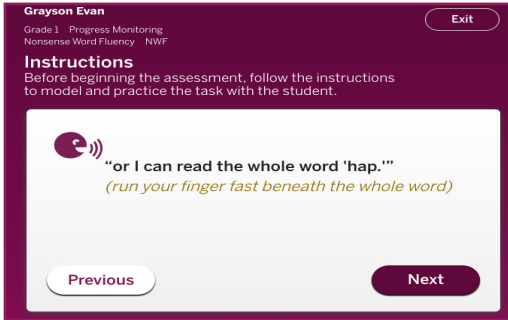
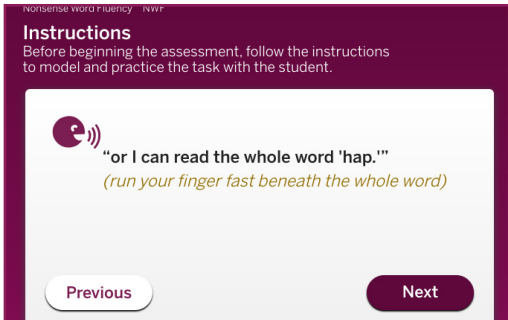
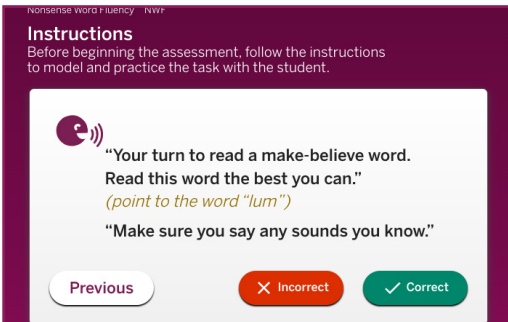
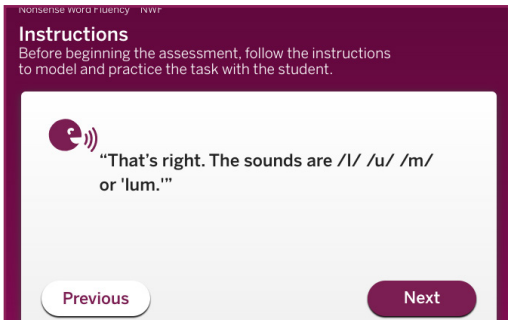
#	Original Directions	Revised Directions
3		"I can read the whole word 'hap'" (run your finger fast beneath the whole word)
4		"Or, if I need to, I can say the sounds of the letters I know, /h/ /a/ /p/." (point to each letter)
5		Your turn to read a make-believe word. Read this word the best you can. But new buttons: Incorrect, Correct/Read whole word, Correct/Sounded out
6		Correct/Read whole word option "That's right. The word is 'lum.'"

Table 1. *Cont'd*

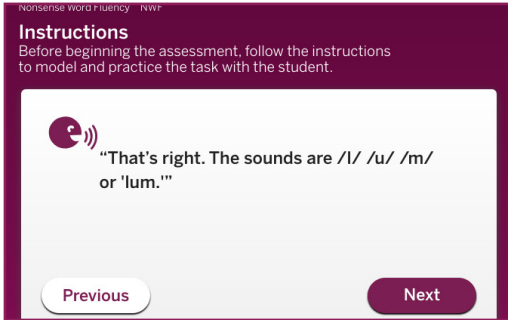
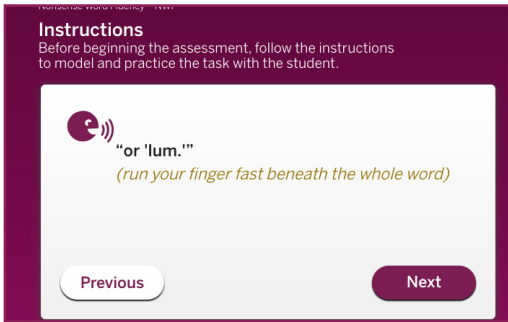
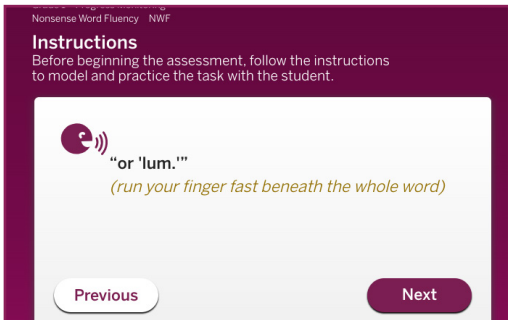
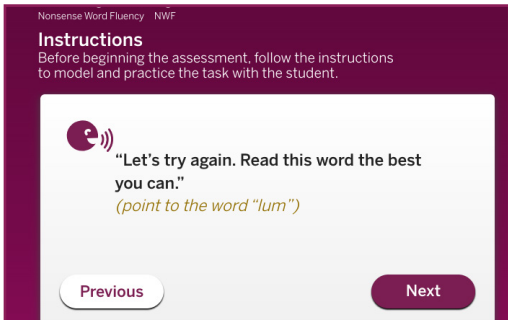
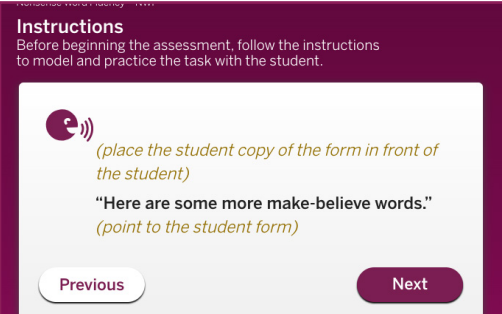
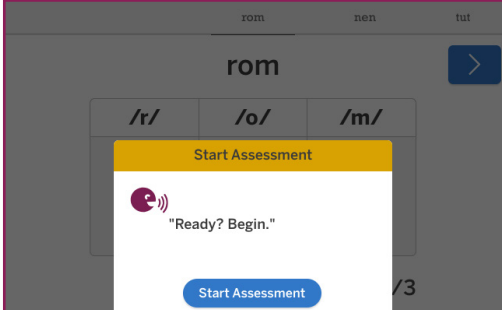
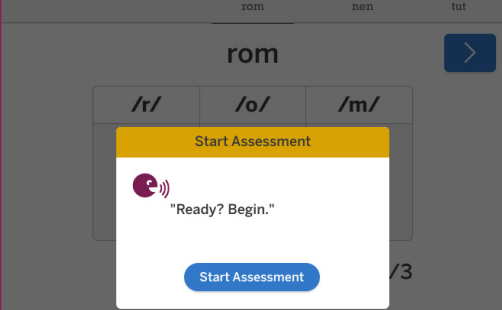
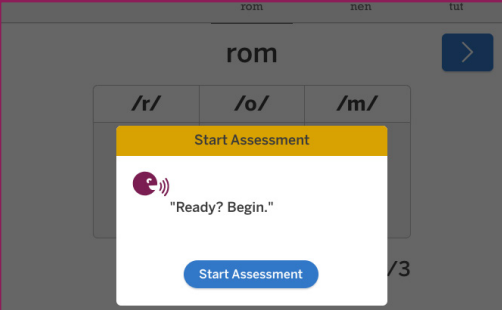
#	Original Directions	Revised Directions
7	 Instructions Before beginning the assessment, follow the instructions to model and practice the task with the student. "That's right. The sounds are /l/ /u/ /m/ or 'lum.'" Previous Next	Correct/Sounded out option That's right. The word is 'lum.' Its sounds are /l/ /u/ /m/."
8	 Instructions Before beginning the assessment, follow the instructions to model and practice the task with the student. "or 'lum.'" <i>(run your finger fast beneath the whole word)</i> Previous Next	Incorrect "Remember, you can say the whole word or, if you need to, you can say the sounds. Watch me: 'lum'" (run your finger fast beneath the whole word)
9	 Instructions Before beginning the assessment, follow the instructions to model and practice the task with the student. "or 'lum.'" <i>(run your finger fast beneath the whole word)</i> Previous Next	"or /l/ /u/ /m/." (point to each letter)
10	 Instructions Before beginning the assessment, follow the instructions to model and practice the task with the student. "Let's try again. Read this word the best you can." <i>(point to the word "lum")</i> Previous Next	"Let's try again. Read this word the best you can." (point to the word "lum")

Table 1. *Cont'd*

#	Original Directions	Revised Directions
11		No change
12		No change
13		“When I say ‘Begin,’ read the words the best you can. Point to the word and read it or, if you need to, point to each letter and tell me the sound. Put your finger on the first word.”
14		No change

METHOD

We used a within-student repeated-observation RCT to examine the effect of NWF administration directions on student scores. Two alternate, grade-level forms of NWF (A and B) were administered to students in Grades K-3 in a single sitting. One administration used the operational standardized administration directions (i.e., control: Original directions), and the other used the revised standardized administration directions (i.e., treatment: Revised directions). The two alternate forms per grade and two different sets of administration directions were counterbalanced using random assignment, resulting in the four assignment scenarios described in [Table 2](#).

Table 2.
Assignment of Nonsense Word Fluency Form and Administration Order

Assignment	1st Administration	2nd Administration
Scenario 1	Form A Original directions	Form B Revised directions
Scenario 2	Form B Original directions	Form A Revised directions
Scenario 3	Form A Revised directions	Form B Original directions
Scenario 4	Form B Revised directions	Form A Original directions

PARTICIPANTS

We recruited a convenience sample of students in Grades K-3 from two sites that had previously administered DIBELS® 8 benchmark assessments and had an existing research relationship with the principal investigators. Both sites were located in Western Oregon. One site was a small, rural charter school. The second site was a summer reading program in a mid-size urban district that included students who required extra

instruction in reading from across seven elementary schools.

All students in Grades K-3 at both sites were invited to participate. Two out of 179 students were excluded because they were absent or because valid assessment scores could not be obtained due to student behavior (e.g., speech disorder and volume). A total of 177 student participants across the two sites and four grades had valid test scores on both NWF administrations and were included in the analysis. [Table 3](#) details the sample demographic characteristics.

Table 3.
Sample Demographic Characteristics (n = 177)

Characteristics	Student Participant Sample	
	n	%
Grade		
K	33	18.8
1	48	27.3
2	44	25.0
3	51	29.0
Gender		
Female	90	51.1
Male	86	48.9
Race/Ethnicity		
African American	1	0.6
American Indian	2	1.1
Asian	1	0.6
Hispanic or Latino	44	25.0
Two or More Races	10	5.7
White	118	67.0
Free or reduced-price lunch		
Missing information	78	44.3
Yes	98	55.7
Special Education Services		
No	134	76.1
Yes	42	23.9
English Language Learner Services		
No	158	89.8
Yes	18	10.2
Form Order		
Original-Revised	92	52.0
Revised-Original	85	48.0
School		
A	78	44.1
B	99	55.9

MEASURES

Dependent Variables

DIBELS® 8 ([University of Oregon, 2018-2023](#)) assesses a range of literacy skills, including decoding, word knowledge, accuracy and fluency with connected text, and comprehension. The subtests offered in each grade are aligned with curriculum and instruction for that grade. All subtests are timed; meaning scores reflect both efficiency and accuracy, and each subtest results in a fluency score that represents the number of units identified correctly in 1 minute.

The current study focused on Nonsense Word Fluency (NWF), a standardized, individually administered measure of a student's ability to decode and read phonetically regular letter combinations that are not actual English words. For example, *tup* is phonetically regular and therefore decodable but is not a real word. The fluency with which students read these nonsense words measures how facile they are with knowledge of letter sounds and letter sound blending. NWF is considered a *pure* measure of decoding because vocabulary and sight word knowledge cannot play a role in recognizing nonsense words. To complete the measure, a student is presented with a page of nonsense words and asked to verbally produce either the whole nonsense word, or the individual letter sounds in the word for one minute, resulting in two scores:

1. *NWF Correct Letter Sounds (NWF-CLS)*: The number of correct letter sounds per minute
2. *NWF Words Read Correctly (NWF-WRC)*: The number of words read or recoded correctly (i.e., words blended correctly) per minute.

The two dependent variables in the study are NWF-CLS and NWF-WRC scores.

Independent Variables

Administration Directions (Original or Revised). Conditional on grade, each student was administered two NWF forms. The main independent variable is administration directions: original or revised. Table 1 details the administration script for the original

and revised directions.

Covariates

Form Order. Form order captures a student's first or second NWF attempt. Given that students are essentially doing the same task twice in one sitting, we want to account for any changes in scores due to the repeated attempt as students may perform better on the second attempt.

Direction Order. Likewise, because students are doing essentially the same task twice, they may not hear the instructions as carefully as the first time. Therefore, we want to test whether the scores are different if they get the instructions in Original-Revised order or Revised-Original order.

Educational Characteristics. We include grade level and school to account for between-unit differences in test scores.

PROCEDURES

Data was collected during the 2024-2026 school year, in May at one site and July at the other site. The assessments were administered by trained data collectors who were students or employees at the University of Oregon. Assessed students were called out of class individually and asked to do some reading activities with the examiner. At one site assessments occurred in small rooms near the main school office. At the other site, assessments occurred in a common area outside of the student's classroom. Each student was administered two alternate NWF forms, randomly ordered, using the two sets of directions, also randomly ordered, in one sitting. The second form was administered immediately following the first form.

ANALYSES

We ran a series of two-level random-intercept linear regression models to estimate the effect of revised directions on the NWF scores. All analyses were conducted using the R programming language ([R Core Team, 2026](#)) and the RStudio interface ([Posit](#)

[Team, 2026](#)). In the first model, we ran a bivariate regression to estimate the effect of form directions (Level 1 independent variable) on the NWF-CLS and NWF-WRC scores, setting original directions as the reference category. In the second model, we adjusted for the effect of second attempt (Level 1 covariate), and accounted for between-school and between-grade differences (Level 2 covariates). In the third model, we also adjusted for directions order (Level 2 covariate), within Original-Revised as the reference category. The equation of the full-model for form f and student i is represented below:

Level 1 (Within student):

$$NWF\ Subscore_{fi} = \beta_{0i} + \beta_{1i} Revised_{fi} + \beta_{2i} Second\ Form_{fi} + \epsilon_{fi}$$

Level 2 (Between student):

$$\beta_{0i} = \gamma_{00} + \gamma_{01} Grade_i + \gamma_{02} Revised/Original\ Direction\ Order_i + \gamma_{03} School\ L_i + \eta_{0i}$$

$$\beta_{1i} = \gamma_{10}$$

$$\beta_{2i} = \gamma_{20}$$

NWF Subscore_{fi} is the NWF score (CLS or WRC) on form f for the i^{th} student. *Revised_{fi}* is a dummy-coded independent variable and equals 1 if the student was given the form with revised directions. β_{1i} is the estimate of interest in all specifications. *Second Form_{fi}* is a dummy-coded covariate and equals 1 for the second form the student attempted. ϵ_{fi} is the Level-1 error term. Level 2 formalizes the difference in the intercept due to between-student factors, including grade, school, and the order of administration directions. Standard errors were clustered at the student-level. Alpha was set a-priori at .05.

RESULTS

[Tables 4-6](#) and [Figure 1](#) summarize the distribution of NWF scores by grade, administration direction, direction order, and form attempt, none of which appear to differ much.

Table 4.

Summary Statistics Across Grade and Administration Direction (n = 177)

Grade	Correct Letter Sounds				Words Read Correct			
	Original		Revised		Original		Revised	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K	28.8	13.3	26.7	14.6	6.2	5.0	5.8	5.1
1	41.8	24.1	42.0	24.0	9.8	8.6	10.1	9.1
2	62.1	34.7	63.0	35.6	15.2	13.4	16.1	13.4
3	100.5	54.3	96.2	51.4	27.7	16.7	27.2	15.8

Table 5.

Summary Statistics Across Grade and Form Attempt (n = 177)

Grade	Correct Letter Sounds				Words Read Correct			
	Attempt 1		Attempt 2		Attempt 1		Attempt 2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K	26.3	13.6	29.2	14.2	6.0	4.7	6.0	5.4
1	41.1	22.9	42.8	25.2	10.1	8.5	9.9	9.2
2	58.7	31.0	66.3	38.5	14.7	12.3	16.6	14.4
3	93.5	48.7	103.2	56.5	26.1	15.0	28.8	17.3

Table 6.

Summary Statistics Across Grade and Direction Order (n = 177)

Grade	Correct Letter Sound				Words Read Correct			
	Original-Revised		Revised-Original		Original-Revised		Revised-Original	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K	27.1	14.0	28.5	13.9	6.2	5.0	5.8	5.2
1	42.8	27.7	41.1	19.9	10.1	10.4	9.9	7.1
2	57.6	33.0	72.0	37.1	13.7	12.5	19.5	14.3
3	87.5	52.8	106.5	51.5	24.8	16.3	29.4	15.9

Figure 1.
Distribution of Nonsense Word Fluency Scores Across Grade and Administration Directions



[Table 7](#) summarizes the results of the multi-level estimate of the effect of administration direction version on NWF-CLS scores. In Model 1, we examined the bivariate relationship. Results indicated that students scored 1.4 points less on NWF-CLS when given revised directions, a nonsignificant ($p = 0.28$) difference. In Model 2, in which we adjusted for the effects of second attempt, grade and school, the inference remained unchanged. On average, there was no effect of revised directions on students' NWF-CLS scores. However, we did find that, on average, students scored 5.8 points higher in their second attempt ($p < 0.001$). In Model 3, we included direction order as an additional covariate. Though the main inference did not change, we found a marginally positive estimate suggesting that when receiving instructions in the Revised-Original order, students scored 7.6 points higher ($p = 0.08$), on average. Across all models, we failed to reject the null hypothesis and found that administration directions did not change NWF-CLS scores, on average. [Figure 2](#) visualizes the effect size of the revised directions on NWF-CLS scores.

Table 7.
Effect of New Administration Directions on NWF-CLS Score

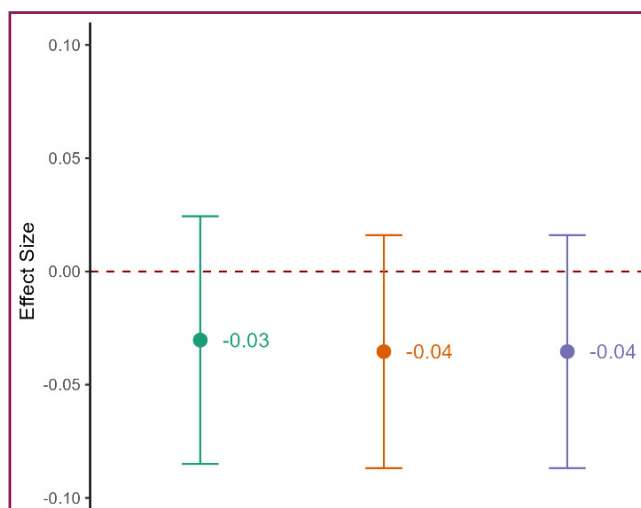
	Model 1	Model 2	Model 3
(Intercept)	61.339*** $p < .001$ $SE = 3.364$	81.267*** $p < .001$ $SE = 4.920$	78.574*** $p < .001$ $SE = 5.129$
Revised	-1.356 $p = .276$ $SE = 1.243$	-1.583 $p = .177$ $SE = 1.169$	-1.583 $p = .177$ $SE = 1.169$
Second Attempt		5.735*** $p < .001$ $SE = 1.169$	5.735*** $p < .001$ $SE = 1.169$
School B		-41.619*** $p < .001$ $SE = 4.403$	-41.422*** $p < .001$ $SE = 4.378$
K		-36.668*** $p < .001$ $SE = 6.610$	-37.755*** $p < .001$ $SE = 6.600$
1		-13.386* $p = .027$ $SE = 6.007$	-14.709* $p = .015$ $SE = 6.020$
3		38.655*** $p < .001$ $SE = 5.911$	36.908*** $p < .001$ $SE = 5.960$
Revised-Original			7.612+ $p = .081$ $SE = 4.357$
SD (Intercept Student)	43.207	27.617	27.441
SD (Observations)	11.691	10.993	10.993
Total observations	354	354	354
R2 Marginal	.000	.563	.568
R2 Conditional	.932	.940	.940
AIC	3336.7	3146.9	3141.1
BIC	3352.2	3181.8	3179.8
ICC	.90	.90	.80
RMSE	2.82	2.80	2.80

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. SE = Standard Error, AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, ICC = Intraclass correlation coefficient, RMSE = Root mean square error.

Figure 2.

Effect Size Estimates of the Effect of Revised Directions on NWF-CLS Scores by Model



[Table 8](#) summarizes results of the multi-level regression analysis for the effect of administration directions on NWF-WRC scores. In Model 1, we examined the bivariate relationship. The results indicated that students scored 0.9 points more on NWF-WRC when given revised directions. However, this difference was not statistically significant ($p = 0.83$). In Model 2, in which we adjusted for effect of second attempt, grade and school, the inference remained unchanged. On average, there was no effect of revised directions on students' NWF-WRC scores. However, we did find that, on average, students scored 1.2 points more on NWF-WRC in their second attempt ($p = 0.003$). In Model 3, we included the direction order as an additional covariate and the inference remained unchanged. In conclusion, we failed to reject the null hypothesis and found that administration directions did not change average NWF-WRC scores. Figure 3 shows the effect size of the revised directions on NWF-WRC scores.

Table 8.
Effect of New Administration Directions on NWF-WRC Score

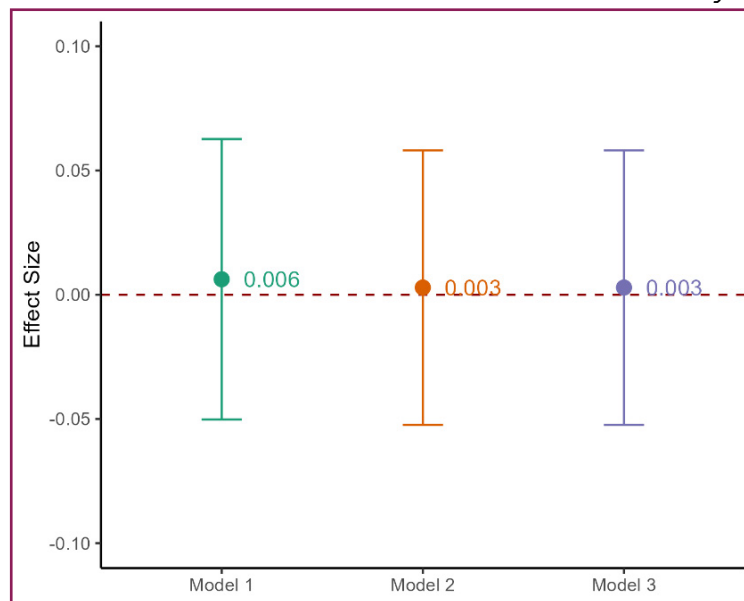
	Model 1	Model 2	Model 3
(Intercept)	15.638*** $p < .001$ $SE = 1.094$	22.112*** $p < .001$ $SE = 1.637$	21.342*** $p < .001$ $SE = 1.710$
Revised	0.090 $p = .829$ $SE = 0.417$	0.042 $p = .919$ $SE = 0.408$	0.042 $p = .919$ $SE = 0.408$
Second Attempt		1.230** $p = .003$ $SE = 0.408$	1.230** $p = .003$ $SE = 0.408$
K		-10.268*** $p < .001$ $SE = 2.195$	-10.578*** $p < .001$ $SE = 2.197$
1		-3.179 $p = .112$ $SE = 1.995$	-3.558+ $p = .077$ $SE = 2.004$
3		12.791*** $p < .001$ $SE = 1.963$	12.291*** $p < .001$ $SE = 1.984$
School B		-14.222*** $p < .001$ $SE = 1.462$	-14.166*** $p < .001$ $SE = 1.457$
Revised-Original			2.176 $p = .134$ $SE = 1.450$
SD (Intercept Student)	14.019	9.135	9.099
SD (Observations)	3.924	3.837	3.837
Total observations	354	354	354
R ² Marginal	.000	.541	.545
R ² Conditional	.927	.931	.931
AIC	2556.7	2399.4	2396.6
BIC	2572.2	2434.2	2435.3
ICC	.90	.90	.80
RMSE	2.82	2.80	2.80

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. SE = Standard Error, AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, ICC = Intraclass correlation coefficient, RMSE = Root mean square error.

Figure 3.

Effect Size Estimates of the Effect of Revised Directions on NWF-WRC Scores by Model



DISCUSSION

Results of this study suggest that revising the administration directions to model the blending approach before sounding out as acceptable behaviors and clarify that the sounding out approach can be used when the reader “needs to” do so did not significantly change students’ scores on the Nonsense Word Fluency (NWF) subtest of the DIBELS® 8 assessment battery. However, the current study has several limitations that should be considered when interpreting the results. First, we recruited a convenience sample of students from two sites in the Pacific Northwest, which limits the generalizability to other populations and regions or schools with different educational contexts. Second, the timing of the data collection also played a role; students may have undergone a full year of schooling and were also exposed to the DIBELS® 8 assessment as part of districts’ assessment cycles, which might have influenced their performance as students were more familiar with the tasks over time. Overall, however, these results support the use of the revised NWF directions, given that they increase face validity without impacting the interpretation of scores relative to those derived from administrations using the original directions.

REFERENCES

- Posit team (2026). *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA. <http://www.posit.co/>
- R Core Team (2026). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- University of Oregon (2018-2023). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS®): Administration and Scoring Guide*, 2023 Edition. Eugene, OR: University of Oregon. Available: <https://dibels.uoregon.edu>