Spanish Instructions as a DIBELS® Accommodation

DIBELS Technical Report 25-S01

University of Oregon (2025). Spanish Instructions as a DIBELS® Accommodation (DIBELS Technical Report 25-S01). Eugene, OR: Author. Available: https://dibels.uoregon.edu

Contributing Authors

Havisha Khurana

Graduate Research Assistant

Department of Education Studies

College of Education, University of Oregon

Gina Biancarosa, Ed. D.

Professor and Ann Swindells Chair in Education

Department of Special Education and Clinical Sciences

College of Education, University of Oregon

Patrick C. Kennedy, Ph.D.

Senior Research Associate and Director of Data Management and Analysis Group

Center on Teaching & Learning

College of Education, University of Oregon

Janet Otterstedt

Senior Research Assistant

Center on Teaching & Learning

College of Education, University of Oregon

Table of Contents

Abstract

An underpowered, randomized control trial was conducted to investigate the impact of administering DIBELS 8th Edition assessment directions in Spanish as an accommodation for Spanish-English bilingual students in kindergarten and Grade 1. Specifically, the study aimed to determine whether providing Spanish directions would improve DIBELS performance, particularly among students with lower English skills. Forty-six kindergarten and first-grade students were randomly assigned to receive directions in English or Spanish. Ordinary least square regression models were used to analyze the effects of language of directions on composite scores, accounting for English proficiency and grade. Results showed a marginal negative effect of giving the directions in Spanish for students with higher English reading skills and marginal positive effect of giving the directions in Spanish for students with lower English reading skills. Results suggest that providing directions in Spanish may be a promising accommodation for Spanish bilingual students with poor English reading skills.

©University of Oregon. All rights reserved.

Spanish Instructions as a DIBELS® Accommodation

DIBELS 8th Edition is a widely used assessment of early literacy skills in kindergarten through eighth grade in the United States. However, standardized procedures require DIBELS to be administered in English, and students are also expected to respond in English. To further develop, validate, and refine DIBELS for young Spanish-English emerging bilingual students, we investigated if the language of assessment directions affects students' performance on DIBELS. As such, we hypothesized that students whose primary language is not English may perform better on DIBELS if the directions are given in their primary language, particularly if their linguistic proficiency in English is limited. Specifically, the study investigated the following research question: Does the performance on DIBELS differ when the language of the directions matches the students' primary language for Spanish-English emerging bilinguals, after adjusting for their English proficiency skills and grade? Hypothesis: If the provision of Spanish language directions is an accommodation as hoped, we hypothesize that students with lower levels of English language skills will benefit more than those with higher levels of English skills. If all students perform better on DIBELS, it suggests that the provision of Spanish language directions is a modification.

Method

We utilized a randomized control trial design to examine the effect of language of directions and modified PSF directions.

Participants

We recruited participants in kindergarten and first grades from three school districts in Oregon. Students meeting two conditions were included in the study: their primary language, as identified by their teacher, was Spanish, and their parent/guardian returned a signed consent letter for participation. Therefore, the study relies on a convenience sample which was then randomly assigned to treatment and control conditions. Forty-six participants across two grades

©University of Oregon. All rights reserved.

were included in the study. Table 1 summarizes the demographic characteristics of the sample.

Measures

Dependent Variables

DIBELS: DIBELS 8 (University of Oregon, 2018-2023) is a series of one-minute assessments administered in kindergarten through grade 8. Subtests administered in kindergarten and first grade include Letter Naming Fluency (LNF), Phonemic Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Word Reading Fluency (WRF). Additionally, Oral Reading Fluency (ORF) is administered in first grade. Each subtest results in a fluency score that represents the number of units identified correctly in one minute. Kindergarten students were administered LNF, PSF, NWF, and WRF, and grade 1 students were administered all five subtests. All participants within a grade received the same forms. In kindergarten, the forms included the beginning-of-year benchmark form for LNF and the progress monitoring form 20 for PSF, NWF and WRF. In first grade, the forms included the beginning-of-year benchmark for LNF, progress monitoring form 20 for PSF, NWF and WRF, and progress monitoring form 16 for ORF. The scores on these subtests were used to calculate a composite score using the end-of-year weighting for LNF in kindergarten (University of Oregon, 2023). Students were given a score of zero if the standard discontinue criteria was met (University of Oregon, 2023). The dependent variable is DIBELS composite score.

Letter Naming Fluency (LNF). This standardized, individually administered task provides a measure of risk relating to future literacy development. Students are presented with a page of upper- and lower-case letters arranged in random order and are asked to name as many letters as they can in one minute. The resulting score is the number of correctly named letters per minute.

Phonemic Segmentation Fluency (PSF). PSF is a standardized, individually administered test of phonemic awareness that assesses a student's ability to segment two- to six-phoneme words into their individual phonemes fluently. To administer PSF, an examiner

©University of Oregon. All rights reserved.

orally presents a series of words and asks students to say the individual phonemes in each word. Forms start with two phonemes, increasing in length to three phonemes in kindergarten and six phonemes in first grade. The examiner continues to present words for one minute. The score is the number of correctly produced phonemes.

Nonsense Word Fluency (NWF). NWF is 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. Students are asked to read a series of nonsense words for one minute, resulting in two scores: the number of Correct Letter Sounds (CLS) per minute, and the number of Words Recoded Correctly (WRC; i.e., words blended correctly) per minute.

Word Reading Fluency (WRF). WRF is a standardized, individually administered test of word reading. Students read as many individual words aloud from a printed word list as they can in one minute. Words are arranged in table format. WRF probes include both sight words (e.g., the, is, school) and regular phonetic words in increasing difficulty within forms and across grades. The resulting score is the number of Words Read Correctly (WRC) per minute.

Oral Reading Fluency (ORF). ORF is a standardized, individually administered test of accuracy and fluency with connected text. The passages are calibrated for the goal level of reading for each grade level. Student performance is measured by having students read a passage aloud for one minute. Words omitted, substituted, and hesitations of more than three seconds are scored as errors. Words self-corrected within three seconds are scored as accurate. The resulting score, totaling the words read correctly and words self-corrected within three seconds, is the number of Words Read Correctly (WRC) per minute. The ratio of WRC to the total number of words attempted gives the ORF-Accuracy score.

DIBELS Composite. The DIBELS composite score is an indicator of a student's

©University of Oregon. All rights reserved.

early reading abilities based on performance across subtests. Scores from various subtests are weighted based on confirmatory factor analyses, accounting for relations among subtests (University of Oregon, 2023). To compute the DIBELS composite score, a zero is imputed for subtests that were not administered due to discontinuation procedures, consistent with DIBELS administration recommendations.

Independent Variables

Language of Directions: Conditional on grade, students were randomly assigned to one of two conditions: English condition (students assessed on DIBELS using standard instructions in English) and Spanish condition (students assessed on DIBELS using instructions in Spanish). This is the independent variable for the analysis.

Covariates

English Language Proficiency Assessment (ELPA). ELPA is used in schools to assess the English skills of students whose primary language is not English. ELPA provides domain scores in the areas of Listening, Reading, Writing, and Speaking. Scores in each domain are divided into five levels with Level 1 corresponding to beginning English skills and Level 5 corresponding to advanced English skills. Results of the ELPA are typically used to determine eligibility for English language instruction services and to inform teachers about their students' skills so they can better meet students' needs. For this research, ELPA scores were requested from school records, and ELPA reading scores were used in the covariate and moderation analysis.

Receptive One-Word Picture Vocabulary Test (ROWPVT-4). The ROWPVT-4 is an untimed, individually administered, norm-referenced assessment of an individual's ability to identify which of the four pictures presented matches a word spoken by the examiner. ROWPVT-4 was used as a concurrent measure of receptive English language skills in the covariate and moderation analysis.

©University of Oregon. All rights reserved.

Procedures

Data for the current study were collected across all school sites during April and May of the 2023-2024 school year. Trained bilingual data collectors pulled students individually from their scheduled classroom activities and engaged them in a brief unscripted rapport-building conversation before starting the assessments. One student was non-responsive and escorted back to class. With all other students, an age-appropriate assent script describing the assessment and study purpose in students' preferred language was presented by the data collector. All students assented, after which, the data collector proceeded with the DIBELS assessment in the assigned language, followed by the ROWPVT in English.

Three data collectors administered the assessments. One of the data collectors was a graduate student in school psychology and had recently completed a course on administering DIBELS assessments. The other two data collectors were trained individually by project staff. Training consisted of first introducing each subtest by reviewing the administration directions and watching an example video. Then the administration directions and scoring procedures were presented in detail, with opportunities to ask questions. Data collectors then scored along with videos of DIBELS administrations. They also practiced administering DIBELS and were provided with feedback. Before administering DIBELS for this research, they were required to achieve an interrater reliability of 90%. When all three data collectors administered DIBELS in the field for the first time, a fourth, experienced DIBELS examiner shadow scored along with them to ensure proper administration procedures were adhered to. After administration, each DIBELS protocol was reviewed to confirm correct scoring and arithmetic.

Analysis Plan

We ran a series of single-level linear regression models[1] to answer the research

1 We have a multi-level data structure, with students nested in schools. For the purpose of this pilot study, we intentionally simplified the analysis by considering single-level models as the primary goal was to get preliminary findings.

©University of Oregon. All rights reserved.

question. All analyses were conducted using the R programming language (R Core Team, 2024) and the RStudio interface (Posit Team, 2024). For analysis, we used the functions from `base` R and `pwr` package (Champely, 2020).

In the first model, we regressed the DIBELS composite score on the language of direction, where directions in English were the reference category. In the second model, we adjusted for students' grades and English skill. In the third model, we tested the interaction between language of direction and students' English skills. We used two proxies to capture students' English skills: Receptive One-Word Picture Vocabulary Test Score (ROWPVT-4) and ELPA reading level. While ROWPVT-4 test score was a continuous measure, we categorized ELPA reading levels so that each category had about the same number of students, resulting in levels 1 & 2, and levels 3, 4 & 5 as the two groups.

Model 1:

$$Composite_i = \beta_0 + \beta_1 + Spanish_i + \epsilon_i$$

$Composite_i$ is the DIBELS 8 composite score for the i^{th} student. $Spanish_i$ is a binary treatment variable and ϵ_i is the student-level error term. β_0 corresponds to the average composite score for students who received English directions. The parameter of interest is β_1 , the mean difference in DIBELS composite score between students receiving Spanish and English directions.

Model 2:

$$Composite_i = \beta_0 + \beta_1 Spanish_i + \beta_2 EnglishSkill_i + \beta_3 Grade_i + \epsilon_i$$

Here, in addition to Model 1, we control for students' grades and students' English skills (cantered scores on the ROWPVT-4 assessment or ELPA categorical levels). The parameter of interest is still β_1, which captures the mean difference in scores for students receiving Spanish and English directions after adjusting for grade-level differences and English skills.

©University of Oregon. All rights reserved.

Model 3:

$$Composite_i = \beta_0 + \beta_1 Spanish_i + \beta_2 EnglishSkill_i + \beta_3 Spanish_i * EnglishSkill_i + \beta_4 Grade_i + \epsilon_i$$

Here, we tested the interaction between language of direction and students' English skills. β_1 represents the mean difference in DIBELS composite score between students receiving Spanish and English directions with an average ROWPVT score or the reference ELPA category. β_3 is the parameter of interest which tests whether students' English skills moderate the effect of Spanish language directions on the DIBELS composite score.

Results

We recruited 46 students across two grades from three school districts for the study. To construct the analytical sample, several research-question specific restrictions were applied to the raw data. First, students who were randomized to the treatment but did not participate in the study were removed (n = 1). We constructed the analytical sample for the research question by removing any students with an invalid administration of one of the required subtests to calculate the composite score (n = 2). Furthermore, students with a missing ROWPVT-4 (n = 2) assessment score or a missing ELPA score (n = 3) were also excluded depending on the covariates used in the analytical approach. Thus, we had between 40 to 41 students in the analytical sample depending on model specification.

Table 2 provides the summary statistics of the analytical sample. About 40% of the students were in kindergarten and the remaining were first-grade students. Of these students, roughly half were given directions in Spanish and English respectively. Fifty-five percent of the students also had a reading level of 1 and 2, and 45% of the students had a reading level of 3 and above. Notably, students who received Spanish directions had a lower average score on the ROWPVT-4 assessment compared to their peers, however, this difference is not statistically significant (t = 1.48, p = 0.15). Additionally, across both model specifications, students who received Spanish directions had lower unadjusted average composite score.

©University of Oregon. All rights reserved.

Table 3 summarizes the results from the regression analysis. In Model 1, we examined the bivariate relationship between the language of directions and composite score. The results indicated that students with average ROWPVT scores who received directions in Spanish scored 14 points lower on the composite score compared to those who received directions in English. However, this difference was not statistically significant ($p = 0.24$), with standard errors of the estimate being as large as the estimate itself. Language of direction explained approximately 3.5% of the variance in composite scores. In Model 2, where we adjusted for grade and ROWPVT-4 scores, the inference remained unchanged. On average, students who received Spanish instructions did not score significantly different than students who received English instructions using an alpha-threshold of 0.05. In Model 3, we included an interaction term to test whether a student's ROWPVT score moderated the effect of the language of directions on composite scores. We found a point estimate of -0.36 for the interaction term, meaning that students with higher English skills benefit less from Spanish directions, but this is not a statistically or practically meaningful difference ($p = 0.53$). In conclusion, our analysis did not find evidence that language of direction affects the DIBELS 8 composite score, on average in a group of multilingual Spanish kindergarten and first-grade students. Additionally, we found no evidence that ROWPVT-4 scores moderate the effect of the language of direction on students' composite scores, on average.

Table 4 presents the regression results with ELPA reading level as the covariate and moderator, with reading level of 3+ as the reference group. In models 1 and 2, we found no effect on directions in Spanish on students' composite scores, on average. In model 3, we tested the heterogeneous effects for directions on Spanish by student reading levels. For students with reading levels of 3 and above, directions in Spanish had a negative effect on composite scores, i.e., students scored 22 points lower on the DIBELS 8 composite scale compared to students receiving English directions ($p = 0.09$), but this difference statistically insignificant at 0.05 alpha-threshold. On the other hand, student with lower ELPA reading levels scored 7 points

©University of Oregon. All rights reserved.

higher on the composite scale when given directions in Spanish compared to students receiving English directions ($p = 0.11$), which was also statistically insignificant at the conventional alpha-threshold levels.

Even though we were aware that the study was under-powered, we ran retrospective power analysis to gauge the sample sizes required to detect significant interaction terms, assuming the population treatment effect is consistent with the effect observed in our analysis, with 80% statistical power. The retrospective power analysis suggested that sample sizes of 969 and 153 would be required to detect a significant difference using the three specification approaches of ROWPVT-4 scores and ELPA reading levels, respectively.

Discussion

The study suggests that providing directions in Spanish on DIBELS 8 assessment as an accommodation had marginal positive impact on student performance for students with lower English reading skills and a marginal negative impact for students with higher English reading skills. However, the current study has several limitations that should be considered when interpreting the results. Though we had balance on key demographic variables that predict student performance, which ensured that randomization was effective, one major limitation was the small sample size, which impacts the precision and the reliability of the estimates. Additionally, we recruited students from Oregon through convenience sampling, which limits the generalization to other populations and regions or schools with different educational contexts. The timing of the data collection also played a role; students may have undergone a full year of schooling in English by the time of the study and were also exposed to the DIBELS 8 assessment as part of districts' progress-monitoring assessment cycles, which might have influenced their performance as students were more familiar with the tasks over time.

Given these factors, future research should aim to recruit larger samples and consider administering assessments at the beginning of the academic year to capture differences before

©University of Oregon. All rights reserved.

students are heavily exposed to the target language or assessment tools and methods.

References

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0, https://CRAN.R-project.org/package=pwr

Posit Team (2024). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/

R Core Team (2024). 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. University of Oregon. https://dibels.uoregon.edu

University of Oregon (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

©University of Oregon. All rights reserved.

Tables

Table 1

Demographic Characteristics of Analytical Sample by Research Question and Model Specification

	ELPA Sample		ROWPVT Sample	
	English	Spanish	English	Spanish
Grade				
Kindergarten	8	8	8	9
1	11	13	11	13
Gender				
Female	8	12	9	12
Male	9	8	8	9
Race				
American Indian/Alaska Native	1	2	1	2
Black/African American	0	2	0	2
Hispanic	13	14	13	15
White	3	2	3	2
Ethnicity				
Hispanic	17	20	17	21
Special Education				
No	14	17	13	18
Yes	3	2	4	2
English Language Learner				
No	1	0	3	1
Yes	18	21	16	21
Free-and-reduced price lunch				
Yes	17	19	17	20
Student Primary Language				
English	3	1	3	1
Spanish	14	19	14	20
Overall ELPA Level				
Emerging	4	6	2	6
Progressing	13	13	13	13
Proficient	2	2	2	2
ELPA Reading Level				
1	5	7	3	7
2	4	6	4	6
3	7	6	7	6
4	1	0	1	0
5	2	2	2	2

Note. Demographic characteristics for 4 students in the analytical sample were incomplete or missing.

©University of Oregon. All rights reserved.

Table 2

Sample Size and Summary Statistics on Key Variables by Research Question and Model Specification

	ROWPVT-4 sample		ELPA sample	
	N	Percent	N	Percent
Grade				
K	17	41.46	16	40.00
1	24	58.54	24	60.00
Language of Directions				
English	19	46.34	19	47.50
Spanish	22	53.66	21	52.50
ELPA Reading Level				
1 & 2			22	55.00
3+			18	45.00
	English	Spanish	English	Spanish
ROWPVT-4 Score	74.53 (16.97)	64.68 (25.43)		
DIBELS 8 Composite Score	431.11 (45.56)	416.32 (34.09)	427.68 (47.40)	416.81 (34.86)

Note. Mean (SD) reported for measures.

©University of Oregon. All rights reserved.

Table 3

Effect of Spanish Directions on DIBELS 8 Composite Score Adjusting for Grade and Receptive Vocabulary.

	Model 1	Model 2	Model 3
DIBELS 8 Composite Score	431.105***	403.967***	402.307***
	(9.131) [<0.001]	(9.465) [<0.001]	(9.886) [<0.001]
Directions in Spanish	-14.787	-8.381	-7.414
	(12.465) [0.243]	(10.299) [0.421]	(10.491) [0.484]
Grade 1		40.490***	40.889***
		(10.113) [<0.001]	(10.213) [<0.001]
ROWPVT-4		0.687**	0.953+
		(0.249) [0.009]	(0.484) [0.057]
Directions in Spanish: ROWPVT-4			-0.363
			(0.567) [0.525]
Num. Obs.	41	41	41
R^2	0.035	0.412	0.418
F	1.407	8.632	6.474

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

Note. Estimate (Std. Error) [p-value] reported in the table. ROWPVT-4 = Receptive One Word Picture Vocabulary Test, Fourth Edition.

©University of Oregon. All rights reserved.

Table 4

Effect of Spanish Directions on DIBELS 8 Composite Score Adjusting for Grade and ELPA Reading Levels.

	Model 1	Model 2	Model 3
DIBELS 8 Composite Score	427.684***	421.902***	427.364***
	(9.469)	(9.285)	(9.642)
	[<0.001]	[<0.001]	[<0.001]
Directions in Spanish	-10.875	-6.533	-22.693+
	(13.068)	(8.875)	(13.021)
	[0.411]	[0.466]	[0.090]
Grade 1		44.457***	47.272***
		(8.952)	(8.902)
		[<0.001]	[<0.001]
ELPA Reading Levels 1 & 2		-42.130***	-57.101***
		(8.903)	(12.516)
		[<0.001]	[<0.001]
Directions in Spanish:ELPA Reading Levels 1 & 2			29.437
			(17.706)
			[0.105]
Num.Obs.	40	40	40
R2	0.018	0.581	0.612
F	0.692	16.630	13.775

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

Note. ELPA Reading Level 3+ is the reference category. Estimate (Std. Error) [p-value] reported in the table. ELPA = English Language Proficiency Assessment.

©University of Oregon. All rights reserved.