



## The Relationship of Oral Reading Fluency Endurance to Comprehension in an Intermediate Depth Orthography

Irene Cadime<sup>a</sup>, Sandra Santos<sup>b</sup>, Fernanda L. Viana<sup>a</sup>, and Iolanda Ribeiro<sup>a</sup>

<sup>a</sup>University of Minho, Braga, Portugal; <sup>b</sup>Center for Research in Higher Education Policies, Matosinhos, Portugal

### ARTICLE INFO

#### Article history:

Received 6 April 2021  
Accepted 16 August 2022  
Available online 18 April 2023

#### Keywords:

Oral reading fluency  
Endurance hypothesis  
Reading comprehension  
Intermediate depth orthography

### ABSTRACT

One-minute oral reading fluency (ORF) tests are widely used, but concerns have been raised regarding whether readers are able to maintain their performance if asked to read for a larger period. The main goals of this study were to investigate whether students are able to maintain their ORF performance in a three-minute task and whether scores measured at one and at three minutes are equally good predictors of the performance in a standardized reading comprehension measure. The sample was composed of 159 Portuguese primary school students (grades 2-4). The results suggested that the number of words read correctly (WRC) declined across reading time and that three-minute mean score underestimated fluency in all grade levels compared to a one-minute reading score. The WRC measured either at one minute or using a three-minute average score was an equal predictor of reading comprehension in all grades. Implications for theory and practice are discussed.

### La relación entre la resistencia en la fluidez lectora y la comprensión en una ortografía de profundidad intermedia

### RESUMEN

Las pruebas de fluidez lectora de un minuto son muy utilizadas, pero se ha planteado la preocupación de si los lectores son capaces de mantener su rendimiento si se les pide que lean durante un período más largo. Los objetivos principales de este estudio han sido investigar si los estudiantes son capaces de mantener su rendimiento de fluidez lectora en una tarea de tres minutos y si las puntuaciones medidas a uno y a tres minutos son igualmente predictoras del rendimiento en una medida estandarizada de comprensión lectora. La muestra estaba compuesta por 159 alumnos portugueses de primaria (2º a 4º curso). Los resultados sugieren que el número de palabras leídas correctamente (PLC) disminuyó a lo largo del tiempo de lectura y que la puntuación media a los tres minutos subestimó la fluidez en todos los cursos en comparación con la puntuación de un minuto de lectura. El PLC medido ya sea en un minuto o utilizando una puntuación media de tres minutos fue un predictor igual de la comprensión lectora en todos los cursos. Se discuten las implicaciones para la teoría y la práctica.

#### Palabras clave:

Fluidez lectora oral  
Hipótesis de resistencia  
Comprensión lectora  
Ortografía de profundidad intermedia

Alphabetic orthographies vary in terms of depth, i.e., in terms of the complexity and consistency of the phoneme-grapheme correspondences (Seymour, 2008). Therefore, they can be classified in a depth continuum, where transparent orthographies (e.g., Finnish) are those that have mainly consistent correspondences between phonemes and graphemes, whereas opaque orthographies (e.g., English) have more complex and inconsistent correspondences. Intermediate depth orthographies, such as European Portuguese, are in the middle of the continuum, being composed of some regularities and some more complex relationships between sounds and letters (Seymour et al., 2003; Sucena et al., 2009). Oral reading fluency

(ORF) can be defined as “the ability to read a text quickly, accurately, and with proper expression” (National Institute of Child Health and Human Development, 2000, p. 3-5). Evidence for a close relationship between ORF and reading comprehension has been systematically provided, not only in opaque (Burns et al., 2011; Chang, 2020; Petscher & Kim, 2011; Wolters et al., 2020), but also in transparent and intermediate depth orthographies (Álvarez-Cañizo et al., 2020; Angelelli et al., 2021; Cadime et al., 2017; Ecalle et al., 2021; Padeliadu & Antoniou, 2014; Recio & León, 2015; Ribeiro et al., 2016).

Curriculum-based measurement (CBM) consists of using the curriculum as a basis for the development of measurement procedures

Cite this article as: Cadime, I., Santos, S., Viana, F. L., & Ribeiro, I. (2023). The relationship of oral reading fluency endurance to comprehension in an intermediate depth orthography. *Psicología Educativa*, 29(2), 143-148. <https://doi.org/10.5093/psed2023a6>

Funding: This study was conducted at the Research Centre on Child Studies (CIEC) and at the Psychology Research Centre (CiPsi), University of Minho, and supported by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education (UID/CED/00317/2019 and UID/PSI/01662/2019), through the national funds (PIDDAC). Correspondence: [ireneCadime@psi.uminho.pt](mailto:ireneCadime@psi.uminho.pt) (I. Cadime).

ISSN: 1135-755X/© 2023 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

to assess the growth in basic skills such as reading (Deno, 1985). Regarding reading fluency, CBM methods usually consist of a grade-level text that children read aloud for one minute, and the number of words read and errors made are recorded and typically combined into one metric: the number of words read correctly per minute (Hosp et al., 2016). This type of assessment has been extensively used (Amendum et al., 2021; Martins & Capellini, 2021; Morris et al., 2017), but some concerns have been raised about its validity. When reading connected text, it is important that students be able to maintain an adequate rate of reading over long periods of time, given that most of the reading tasks, in and out of school, require the reading of texts with some extension that cannot possibly be fully read in only one minute. According to Deeney (2010), on the one hand, it is possible that this ability is impaired in children struggling with reading, who may not have the necessary endurance for maintaining their initial reading accuracy and speed, so their performance will decline with time. Deeney (2010) also argues that, on the other hand, it is possible that some other students have more difficulty during the first minute but increase their accuracy and rate over time due to an improvement in the comprehension of the text content and the consequent mobilization of semantic cues in the reading. In either case, one-minute tests would be inadequate for assessing students.

Some evidence for this “endurance hypothesis”, i.e., that students’ performance declines with the time of reading, was firstly provided in a report performed in the context of the National Assessment of Education Progress, where a sample of fourth graders was studied (Daane et al., 2005). Students read a text composed of 198 words and were not instructed to read the passage quickly, so it was assumed that the reading rate reflected the natural pace of the students. Two ORF indices were calculated for each student: (a) number of correct words read “during” the first minute of reading and (b) “average” number of words read correctly per minute (WRCPM) based on the full duration of each student’s reading. Their results indicated that, if only the first minute of reading was considered, 56% of the sample read 130 words or more, whereas if the average number of words read correctly per minute was considered, only 38% of the sample read 130 words or more.

Some additional evidence concerning this loss of endurance was presented by Valencia et al. (2010). They compared the mean scores in accuracy (measured by the percentage of errors in reading), the rate of reading and the WRCPM after one and after three minutes of reading using a sample composed of students in second, fourth, and sixth grades. The results for the second and fourth graders showed a statistically significant decrease in the mean accuracy, rate of reading, and WRCPM scores when students read for three minutes compared with the scores obtained after only one minute of reading. This result indicates that students lose accuracy and speed if asked to read for longer intervals of time. In the case of sixth graders, a significant decrease was only observed for accuracy scores.

A related question is whether scores obtained in longer measures of reading fluency are better indicators of comprehension than shorter measures. Research with primary school students has consistently shown that one-minute measures, such as DIBELS, provide scores that are strong predictors of reading comprehension, not only in English (Espin & Deno, 2016; Morris et al., 2017; Reschly et al., 2009), but also in more transparent orthographies (Massonnié et al., 2019). Research on the relationship between these two variables using longer ORF measures is not so abundant, but some studies have shown medium to high correlations between these scores and reading comprehension measures in a wide range of orthographies in primary school years (Angelelli et al., 2021; Nevo et al., 2020; Santos et al., 2017). However, a direct comparison of the findings of these studies is not possible due to variations in the reading comprehension measures, type, and length of texts used to assess fluency, and the range of grades covered in the study. To our knowledge, there are only a couple of studies comparing conjointly the predictive power

of the scores in one-minute ORF tests and in longer measures. In the study conducted by Daane et al. (2005) described above, the results showed that both scores (WRCPM at 1-minute or a 3-minute mean) had positive correlations with reading comprehension. Valencia et al. (2010) also examined correlations between reading comprehension scores and the three ORF indicators measured at the two time intervals. They found that correlations between comprehension and the rate and WRCPM scores were similar regardless of whether ORF scores were measured after one or three minutes of time. Regarding accuracy, although the authors indicated that “correlations between accuracy and ITBS comprehension increased at 3 minutes of reading for grades 2 and 4, suggesting that, in general, having students read for slightly longer may increase the predictive power of accuracy scores” (Valencia et al., 2010, p. 278), differences were not meaningful (grade 2:  $-.37$  vs.  $-.43$ ; grade 4:  $-.35$  vs.  $-.42$ ).

In conclusion, although there is evidence that there is a decrease in ORF scores in longer tasks, compared to one-minute tasks, the association with reading comprehension seems to be similar. However, as indicated before, only two studies investigated the relationship between reading comprehension and ORF measured at two time intervals. Historically, the CBM-based score WRCPM has been used as a measure of general reading proficiency and therefore this claim can be only made if validity evidence is presented, namely validity evidence based on the relations to other variables such as reading comprehension (Espin & Deno, 2016). Therefore, more studies are needed to address whether ORF scores measured at one minute or at longer times are equally valid. Moreover, the research described was conducted with English readers and, therefore, it is unclear whether results can be generalized to other languages with a different degree of orthographic depth. As previously indicated, English is an opaque orthography with a very high number of complexities and inconsistencies between sounds and letters (Seymour, 2008; Seymour et al., 2003). Therefore reading automaticity is harder to develop in English than in more transparent orthographies (Suggate et al., 2014). Thus, it remains unclear whether performance decreases across time observed in longer ORF tasks also occurs when reading in intermediate and transparent orthographies, where automaticity is developed much easier.

The main goal of the present study was to explore the ORF endurance hypothesis (Deeney, 2010; Valencia et al., 2010). To achieve this goal, we investigated whether the three commonly used indicators of ORF – accuracy (number of errors during reading), speed (rate of reading), and an indicator that combines both accuracy and speed of reading (WRCPM) – vary as a function of the time interval during which they are measured and if they are equally good predictors of reading comprehension in an intermediate depth orthography – European Portuguese. The following research questions were addressed: (1) does reading accuracy and speed decrease with reading time?; (2) how do informal ORF scores at one and three minutes predict reading comprehension on a standardized comprehension measure? Based on the results of previous studies, conducted with English readers, we expect a decrease in accuracy and speed with reading time. We also expect that ORF scores taken at one- and three-minute reading are equally strong predictors of reading comprehension.

## Method

### Participants

The sample consisted of 159 students from the second ( $n = 50$ , mean age = 7.88,  $SD = 0.33$ ; 64% were girls), third ( $n = 52$ , mean age = 9.06,  $SD = 0.37$ ; 44.2% were girls), and fourth grades ( $n = 57$ , mean age = 10.16,  $SD = 0.41$ ; 50.9% were girls). All students attended public schools in the north of Portugal. The sample included only children

who had European Portuguese as their first language. None of the children were receiving special education services.

## Measures

### Oral Reading Fluency

Three narrative texts were used to measure oral reading fluency: *Pirate's son* (706 words), *A Cat with special powers* (667 words), and *Christmas in Boticas* (1161 words). The first two texts were the transcriptions of a listening comprehension test (Santos et al., 2015) and the third was retrieved from an unpublished ORF test created by the research team. Students were asked to read each text aloud, with accuracy, speed, and with proper expression, and the test administrator monitored the duration of the reading and wrote down students' errors. The order of texts' administration was counterbalanced. Students' reading was also recorded for later rescoring and interrater reliability checks.

After the assessment, three ORF scores were computed for each text: (1) accuracy – scored as the number of errors in each minute. Mispronunciations, substitutions, insertions, omissions, and words incorrectly read and not corrected within 3 seconds were scored as errors. Self-corrections within 3 seconds after the error, repeated words, mispronunciations due to dialect or regional variations, hesitations, or words read slowly but correctly were not scored as errors; (2) rate – scored as the number of words read, correctly or incorrectly; (3) WRC – number of words read correctly. The three types of scores were calculated at 1 minute, 2 minutes and 3 minutes of reading. The mean number of errors and WRC per minute (WRCPM) considering the total three minutes were also computed.

Fifteen students (five from each grade) were randomly selected from the database and their readings were independently scored by a second rater. Interrater agreement, assessed using the intraclass correlation coefficient, ranged between .99 and 1.00.

### Reading Comprehension

To measure reading comprehension, we used a standardized test – the Reading Comprehension test with narrative texts (Santos et al., 2017; Santos et al., 2016) – validated for Portuguese primary school students. This is a norm-referenced test that is composed of three vertically scaled test forms (TRC-n-2, TRC-n-3, and TRC-n-4), each aimed at assessing second, third, and fourth graders, respectively. Each test form includes four texts to be read silently by the students and 27 multiple choice items with three options that assess literal comprehension, inferential comprehension, critical comprehension, and reorganization. Each correct answer is scored as 1. Reliability coefficients (Kuder-Richardson formula 20, person separation reliability, and item separation reliability) of the test forms ranged between .70 and .96.

## Procedure

Legal authorizations for data collection were obtained from the Portuguese Ministry of Education, school boards, and parents or legal tutors. Only students who returned the signed consent form were assessed. The ORF test was administered to students individually. The reading comprehension test was administered collectively to students in a single session in their classroom, without time limit. All tests were administered by trained psychologists, who followed the standardized instructions from the test manuals.

## Statistical Analyses

In a first set of analyses, we considered the number of errors and WRC measured during the first minute of reading (minute 1),

between the first and the second minute (minute 2), and between the second and the third minute (minute 3) obtained in the ORF measure. Rate was excluded from the analyses because exploratory analysis showed that the correlation between rate and WRC at each minute was almost perfect, i.e., close to one. In order to assess the associations between time (i.e., minutes of reading) and the number of errors and WRCPM, two linear mixed models were calculated (one for each ORF indicator). In all models the fixed effects were time and grade, and the random effects were individuals and texts (responses from the same individual are correlated, as well as readings from the same text). The lme4 package (Bates et al., 2015) for R statistical software was used for this analysis.

In a second set of analyses, the mean number of errors and WRCPM considering the total three minutes were used. For this second set of analyses the scores obtained in the three texts were averaged. Paired-samples *t*-tests were computed to compare the mean of the three minutes with the performance obtained in the first minute. Linear regression models were computed to investigate the contribution of these ORF scores to reading comprehension. This second set of analyses was performed using IBM® SPSS Statistics 27.

## Results

Table 1 presents the results of linear mixed models. Regarding the number of errors, the results of the linear mixed model indicated a significant effect of time ( $B = 0.157$ ,  $SE = 0.057$ ,  $p < .001$ ). For each minute more, in average, the number of errors increase in 0.157 words. This effect occurs irrespective of the grade level, given that the interaction effect time  $\times$  grade was not significant ( $p > .05$ ).

Regarding the number of words read correctly (WRC) in each minute and rate, the main effect of time was also significant ( $B = -6.260$ ,  $SE = 0.282$ ,  $p < .001$  and  $B = -6.175$ ,  $SE = 0.271$ ,  $p < .001$ , respectively). Per each minute more, in average, the number of words read correctly decreases 6.260 words and the rate of reading decreases 6.175 words. For these two scores, the interaction effect time  $\times$  grade was not significant ( $p > .05$ ), that is, the effect of time occurs irrespective of the grade level (see Table 1).

**Table 1.** Results for the Linear Mixed Models to Test for the Effects of Minute of Reading and Grade in Each Oral Reading Fluency Indicator

	WRC			Number of errors		
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>B</i>	<i>SE</i>	<i>p</i>
<b>Fixed effects</b>						
(Intercept)	91.351	3.268	< .001	2.302	0.230	< .001
Time	-6.260	0.282	< .001	0.157	0.057	< .001
<b>Grade</b>						
2 (ref.)						
3	16.608	4.508	< .001	-0.231	0.280	0.410
4	32.036	4.410	< .001	-0.069	0.273	0.801
<b>Random effects</b>						
ICC individual	0.895			0.310		
Between-individual variation	659.22			1.52		
ICC text	0.002			0.054		
Between-text variation	1.50			0.27		

Note. WRC = words read correctly; Ref. = reference group; *B* = estimates; *SE* = standard error; ICC = intraclass correlation coefficient.

Table 2 presents descriptive statistics and paired-samples *t*-tests of differences in the number of errors and WRCPM considering either one minute of reading or the mean value obtained in three minutes of reading. The results indicate that the error score does not vary significantly when considering only the number of errors during the first minute of reading or the mean per minute over the three minutes. However, there were significant differences in WRCPM,

as it was higher if only the first minute of reading was considered compared to a three-minute mean score.

**Table 2.** Descriptive Statistics and Paired-Samples *t*-tests for the Errors and Number of Words Read Correctly per Minute in Each Grade

Variable	Mean	SD	Min.	Max.	<i>t</i> (df)	<i>p</i>
<b>Grade 2</b>						
Errors (1 minute)	2.87	1.31	0.67	6.67	0.619(49)	.539
Errors (mean 3 minutes)	2.80	1.37	0.67	6.78		
WRCPM (1 minute)	84.67	19.04	39.67	132.00	11.507(49)	< .001
WRCPM (mean 3 minutes)	78.96	17.70	40.11	122.00		
Reading comprehension	16.84	3.65	9.00	24.00		
<b>Grade 3</b>						
Errors (1 minute)	2.48	1.58	0.33	7.33	-0.727(51)	.470
Errors (mean 3 minutes)	2.57	1.55	0.11	7.11		
WRCPM (1 minute)	104.71	24.57	67.00	175.00	14.789(51)	< .001
WRCPM (mean 3 minutes)	97.94	23.94	60.44	168.00		
Reading comprehension	16.92	4.52	8.00	25.00		
<b>Grade 4</b>						
Errors (1 minute)	2.65	1.57	0.00	6.67	-0.954(56)	.344
Errors (mean 3 minutes)	2.73	1.45	0.67	6.44		
WRCPM (1 minute)	122.37	28.93	68.00	173.00	19.076(56)	< .001
WRCPM (mean 3 minutes)	114.11	27.95	61.22	161.00		
Reading comprehension	17.56	4.41	6.00	26.00		

Note. WRCPM = words read correctly per minute; SD = standard deviation; Min. = minimum; Max. = maximum; *df* = degrees of freedom.

The number of reading errors did not predict reading comprehension for the second grade but had a significant and negative relationship with reading comprehension in the remaining grades, regardless of the time used to compute the scores (see Table 3). WRCPM was a significant reading comprehension predictor in all grade levels, regardless of the considered time spent reading (see Table 3). Moreover, the percentage of variance explained by a one-minute score and a three-minute mean score was similar.

**Table 3.** Regression Models of the Number of Errors and Words Read Correctly per Minute (1 or 3 minutes of reading) Predicting Reading Comprehension Scores in Each Grade

	$\beta$	<i>p</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> adj.
<b>Grade 2</b>				
Errors (1 minute)	.113	.441	.013	-.008
Errors (mean 3 minutes)	.029	.842	.001	-.020
WRCPM (1 minute)	.491	< .001	.241	.225
WRCPM (mean 3 minutes)	.498	< .001	.248	.232
<b>Grade 3</b>				
Errors (1 minute)	-.374	.006	.140	.123
Errors (mean 3 minutes)	-.456	.001	.208	.192
WRCPM (1 minute)	.501	< .001	.251	.236
WRCPM (mean 3 minutes)	.521	< .001	.272	.257
<b>Grade 4</b>				
Errors (1 minute)	-.515	< .001	.265	.252
Errors (mean 3 minutes)	-.496	< .001	.246	.232
WRCPM (1 minute)	.613	< .001	.376	.364
WRCPM (mean 3 minutes)	.610	< .001	.372	.361

Note. WRCPM = words read correctly per minute.

### Discussion

The main goal of this study was to test the so-called ORF “endurance hypothesis” in an intermediate depth orthography. The first research question addressed whether ORF decreased across

the reading time. The results indicate that the number of errors increased, and the WRC decreased as students advanced in reading in all grade levels. Consequently, in all grade levels, the number of words read during the first minute is higher than the average number of words read per minute as computed from the three-minute reading. This finding is similar to those obtained in studies conducted with English readers (Daane et al., 2005; Valencia et al., 2010). Therefore, the loss of endurance in reading seems to occur in all grade levels and is independent of the orthography, i.e., it does not seem to be exclusive of orthographies with more irregularities, in which reading automatization takes longer. Rather than due to linguistic factors, the observed changes in reading across time might be related to readers’ attentional mechanisms. Research has shown that the generation of the phonological code from print is not modular and that it requires attentional processes which are critical for reading (Reynolds & Besner, 2006; Shaywitz & Shaywitz, 2008). Thus, fluency tasks require sustained attention, which “comprises the vigilance and persistence necessary for maintaining concentration during continuous, repetitive activity” (Chenault et al., 2006, p. 245). Therefore, it is likely that the longer the reading task, the more difficult for the reader to maintain his performance, leading to a higher number of errors and to a lower speed. Future studies should address this hypothesis by including sustained attention measures when addressing the performance across time in ORF measures.

The second research question explored whether informal ORF scores measured during the first minute were more predictive of reading comprehension, as measured by a standardized test, than scores calculated taking the three-minute reading. On the one hand, the number of errors during reading was a significant predictor of reading comprehension in grades 3 and 4, regardless of whether one or three minutes of reading were considered. The reasons for the inexistence of a significant effect in grade 2 are unclear but might be related to a low decoding efficiency in this grade, which increases in the following grades. Additionally, research in European Portuguese has shown that, in this orthography, reading speed is a stronger predictor of reading comprehension than accuracy (e.g., Cadime et al., 2017; Santos et al., 2020). On the other hand, the WRCPM was a significant comprehension predictor in all grade levels, regardless of the time spent reading used to compute WRCPM, and the size of the relationship between the ORF indicators and comprehension did not vary whether using a one-minute or a three-minute mean score. These findings are similar to those found in studies in a more opaque orthography – English (Daane et al., 2005; Valencia et al., 2010). More notable is that these results are similar regardless of some methodological differences between studies. For example, in the studies by Valencia et al. (2010) and Daane et al. (2005), the tasks used to measure ORF also included comprehension questions, so that students knew that the goal of the reading task was to comprehend the texts, and instructions did not include any reference to the necessity of reading fast. On the contrary, in our study instructions stated that the goal was to read quickly, without errors and with appropriate prosody, and no comprehension questions about these texts were presented to the students. Thus, the decrease in performance in ORF tasks across time seems to be independent of the goal of the task perceived by the students. A second difference between our study and the one conducted by Valencia et al. (2010) was that they used different texts in different grades to assess ORF, but in our study the same three texts were used in all grade levels. Additionally, they used narrative and expository texts to measure ORF and the standardized test they used to measure comprehension included several text genres, whereas only narrative texts were used in our study. This variation in text genres included not only in the ORF task but also in the standardized reading comprehension measure also does not seem to have an impact on the findings that ORF and reading comprehension are positively associated and that the size of this relationship does not vary as a function of the duration of the ORF measurement.



Thus, our results suggest that scores obtained in one-minute ORF tests are as valid as scores obtained in three-minute tests, whether considering accuracy or speed ORF scores. This finding has important practical implications for assessment in educational settings: if both scores are equally valid, one-minute ORF tests should be preferred, for a more parsimonious assessment of readers. Using shorter measures is particularly important, given that, in a CBM system, students are measured frequently over time and scores are used to evaluate the effects of instruction on their progress (Espin & Deno, 2016). If the scores reveal a lack of progress, then changes in the instruction must be introduced (Bruhn et al., 2019). Thus, with a high number of students per class and frequent measurements that, in a CBM system, are usually weekly (Gesel & Lemons, 2020; Watt et al., 2020), using one-minute ORF measures instead of longer measures reduces significantly the time allocated for assessment and leaves more time for the instruction.

A limitation of our study is related to the use of a small and not representative sample, which limits the possibility of generalization of findings. Future replication studies should include a more representative sample. A second limitation was that, given the low sample size in each grade level, it was not possible to study separately the performance of children with reading difficulties, given the loss of statistical power. As Deeney (2010) states, the lack of endurance – the ability to persist in the reading task across extended periods of time – may be particularly substantial in struggling readers or children with reading difficulties. Therefore, the decrease in performance in longer ORF tasks is expected to be more severe and there is a possibility that the one-minute tasks underrepresent the reading comprehension deficits of these students, contrarily to what happens with typical readers. Future studies with children with reading disabilities are also crucial because CBM methods, that include one-minute ORF measures, have systematically been used in response-to-intervention approaches as tools to identify students with reading disabilities and to make placement decisions for these students (Espin & Deno, 2016). Hence, validity evidence of ORF scores measured at different time intervals to achieve these purposes is also needed.

Another limitation of this study was that prosody was not analyzed. Prosody has been found as one of the main predictors of reading comprehension, especially in more advanced grades, when decoding is already mastered (Fernandes et al., 2018; Wolters et al., 2020). Reading with prosody requires that the reader has segmented text according to major syntactic and semantic elements, which is crucial for the comprehension of the written text (Kuhn & Stahl, 2003). Although we found a decrease in speed and accuracy across the time of reading, it is unclear if this decrease is accompanied or not by a decrease in prosody. It is also unclear whether the prosody levels collected in one-minute measures underrepresent readers' construction of meaning from text or, on the contrary, whether a decrease in prosody as the reading advances leads to a less proficient use of information about syntactic and semantic boundaries and, consequently, to a lower comprehension level. Thus, future studies should not only consider the number of errors, rate of reading and WRCPM, but also prosody indicators.

In conclusion, the results of this study extend the findings of previous research that found evidence for a loss of endurance across reading time in ORF tasks by showing that this effect is also verified in orthographies other than English, but also emphasize that scores derived from one- or three-minutes of reading are similar indicators of reading comprehension and can be both be taken as indicators of general reading proficiency.

## References

Álvarez-Cañizo, M., Cueva, E., Cuetos, F., & Suárez-Coalla, P. (2020). Reading fluency and reading comprehension in Spanish secondary students. *Psicothema*, 32(1), 75-83. <https://doi.org/10.7334/psicothema2019.196>

- Amendum, S. J., Conradi Smith, K., & Liebfreund, M. D. (2021). Explaining reading variance by student subgroup: should we move beyond oral reading fluency? *Journal of Research in Reading*, 44(4), 757-786. <https://doi.org/10.1111/1467-9817.12371>
- Angelelli, P., Romano, D. L., Marinelli, C. V., Macchitella, L., & Zoccolotti, P. (2021). The simple view of reading in children acquiring a regular orthography (Italian): A network analysis approach. *Frontiers in Psychology*, 12, Article 686914. <https://doi.org/10.3389/fpsyg.2021.686914>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Bruhn, A. L., Wehby, J. H., & Hasselbring, T. S. (2019). Data-based decision making for social behavior: Setting a research agenda. *Journal of Positive Behavior Interventions*, 22(2), 116-126. <https://doi.org/10.1177/1098300719876098>
- Burns, M. K., Kwoka, H., Lim, B., Crone, M., Haegele, K., Parker, D. C., Petersen, S., & Scholin, S. E. (2011). Minimum reading fluency necessary for comprehension among second-grade students. *Psychology in the Schools*, 48(2), 124-132. <https://doi.org/10.1002/pits>
- Cadime, I., Rodrigues, B., Santos, S., Viana, F. L., Chaves-Sousa, S., do Céu Cosme, M., & Ribeiro, I. (2017). The role of word recognition, oral reading fluency and listening comprehension in the simple view of reading: a study in an intermediate depth orthography. *Reading and Writing*, 30(3), 591-611. <https://doi.org/10.1007/s11145-016-9691-3>
- Chang, I. (2020). Influences of executive function, language comprehension, and fluency on young children's reading comprehension. *Journal of Early Childhood Research*, 18(1), 44-57. <https://doi.org/10.1177/1476718X19875768>
- Chenault, B., Thomson, J., Abbott, R. D., & Berninger, V. W. (2006). Effects of prior attention training on child dyslexics' response to composition instruction. *Developmental Neuropsychology*, 29(1), 243-260. [https://doi.org/10.1207/s15326942dn2901\\_12](https://doi.org/10.1207/s15326942dn2901_12)
- Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Orange, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading*. U.S. Department of Education, Institute of Education Sciences.
- Deeney, T. A. (2010). One-minute fluency measures: Mixed messages in assessment and instruction. *The Reading Teacher*, 63(6), 440-450. <https://doi.org/10.2307/25615834>
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219-232. <https://doi.org/10.1177/00144029850200303>
- Ecalte, J., Dujardin, E., Gomes, C., Cros, L., & Magnan, A. (2021). Decoding, fluency and reading comprehension: Examining the nature of their relationships in a large-scale study with first graders. *Reading and Writing Quarterly*, 37(5), 444-461. <https://doi.org/10.1080/10573569.2020.1846007>
- Espin, C. A., & Deno, S. L. (2016). Conclusion: Oral reading fluency or reading aloud from text: An analysis through a unified view of construct validity. In K. D. Cummings & Y. Petscher (Eds.), *The fluency construct: Curriculum-based measurement concepts and applications* (pp. 365-384). Springer. <https://doi.org/10.1007/978-1-4939-2803-3>
- Fernandes, S., Querido, L., Verhaeghe, A., & Araújo, L. (2018). What is the relationship between reading prosody and reading comprehension in European Portuguese? Evidence from grades 2 to 5. *Journal of Research in Reading*, 41(S1), S102-S129. <https://doi.org/10.1111/1467-9817.12248>
- Gesel, S. A., & Lemons, C. J. (2020). Comparing schedules of progress monitoring using curriculum-based measurement in reading: A replication study. *Exceptional Children*, 87(1), 92-112. <https://doi.org/10.1177/0014402920924845>
- Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). *The ABCs of CBM: A practical guide to curriculum-based measurement* (2nd ed.). Guilford.
- Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95(1), 3-21. <https://doi.org/10.1037/0022-0663.95.1.3>
- Martins, M. A., & Capellini, S. A. (2021). Identification of struggling readers or at risk of reading difficulties with one-minute fluency measures. *Psicologia: Reflexão e Crítica*, 34(10), 1-9. <https://doi.org/10.1186/s41155-021-00174-z>
- Massonnié, J., Bianco, M., Lima, L., & Bressoux, P. (2019). Longitudinal predictors of reading comprehension in French at first grade: Unpacking the oral comprehension component of the simple view. *Learning and Instruction*, 60, 166-179. <https://doi.org/10.1016/j.learninstruc.2018.01.005>
- Morris, D., Trathen, W., Perney, J., Gill, T., Schlagal, R., Ward, D., & Frye, E. M. (2017). Three DIBELS tasks vs. three informal reading/spelling tasks: A comparison of predictive validity. *Reading Psychology*, 38(3), 289-320. <https://doi.org/10.1080/02702711.2016.1263700>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction*. U.S. Government Printing Office.
- Nevo, E., Vaknin-Nusbaum, V., Brande, S., & Gambrell, L. (2020). Oral reading fluency, reading motivation and reading comprehension

- among second graders. *Reading and Writing*, 33(8), 1945-1970. <https://doi.org/10.1007/s11145-020-10025-5>
- Padeliadu, S., & Antoniou, F. (2014). The relationship between reading comprehension, decoding, and fluency in Greek: A cross-sectional study. *Reading & Writing Quarterly*, 30(1), 1-31. <https://doi.org/10.1080/10573569.2013.758932>
- Petscher, Y., & Kim, Y. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology*, 49(1), 107-129. <https://doi.org/10.1016/j.jsp.2010.09.004>
- Recio, P., & León, J. A. (2015). La lectura en un contexto bilingüe: fluidez y comprensión lectora en alumnos de 1º y 4º de primaria. *Psicología Educativa*, 21(1), 47-53. <https://doi.org/10.1016/j.pse.2015.03.002>
- Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, 47(6), 427-469. <https://doi.org/10.1016/j.jsp.2009.07.001>
- Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance*, 32(6), 1303-1323. <https://doi.org/10.1037/0096-1523.32.6.1303>
- Ribeiro, I., Cadime, I., Freitas, T., & Viana, F. L. (2016). Beyond word recognition, fluency, and vocabulary: The influence of reasoning on reading comprehension. *Australian Journal of Psychology*, 68(2), 107-115. <https://doi.org/10.1111/ajpy.12095>
- Santos, S., Cadime, I., Viana, F. L., Chaves-Sousa, S., Gayo, E., Maia, J., & Ribeiro, I. (2017). Assessing reading comprehension with narrative and expository texts: Dimensionality and relationship with fluency, vocabulary and memory. *Scandinavian Journal of Psychology*, 58(1), 1-8. <https://doi.org/10.1111/sjop.12335>
- Santos, S., Cadime, I., Viana, F. L., Prieto, G., Chaves-Sousa, S., Spinillo, A. G., & Ribeiro, I. (2016). An application of the Rasch model to reading comprehension measurement. *Psicología: Reflexão e Crítica*, 29(38), 1-19. <https://doi.org/10.1186/s41155-016-0044-6>
- Santos, S., Cadime, I., Viana, F. L., & Ribeiro, I. (2020). Cross-lagged relations among linguistic skills in European Portuguese: A longitudinal study. *Reading Research Quarterly*, 55(2), 177-192. <https://doi.org/10.1002/rrq.261>
- Santos, S., Viana, F. L., Ribeiro, I., Prieto, G., Brandão, S., & Cadime, I. (2015). Development of listening comprehension tests with narrative and expository texts for Portuguese students. *The Spanish Journal of Psychology*, 18(e5), 1-7. <https://doi.org/10.1017/sjp.2015.7>
- Seymour, P. (2008). Early reading development in European orthographies. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 296-315). Blackwell Publishers Ltd.
- Seymour, P., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143-174. <https://doi.org/10.1348/000712603321661859>
- Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: The neurobiology of reading and dyslexia. *Development and Psychopathology*, 20(4), 1329-1349. <https://doi.org/10.1017/S0954579408000631>
- Sucena, A., Castro, S. L., & Seymour, P. (2009). Developmental dyslexia in an orthography of intermediate depth: The case of European Portuguese. *Reading and Writing*, 22(7), 791-810. <https://doi.org/10.1007/s11145-008-9156-4>
- Suggate, S., Reese, E., Lenhard, W., & Schneider, W. (2014). The relative contributions of vocabulary, decoding, and phonemic awareness to word reading in English versus German. *Reading and Writing*, 27(8), 1395-1412. <https://doi.org/10.1007/s11145-014-9498-z>
- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion and consequential validity. *Reading Research Quarterly*, 45(3), 270-291. <https://doi.org/10.1598/RRQ.45.3.1>
- Watt, S., Conoyer, S., Ford, J. W., Foegen, A., & Luckey, G. (2020). Raising the power of curriculum-based measurement tools in preservice training. *Teacher Education and Special Education*, 44(1), 78-92. <https://doi.org/10.1177/0888406420916829>
- Wolters, A. P., Kim, Y. S. G., & Szura, J. W. (2022). Is reading prosody related to reading comprehension? A meta-analysis. *Scientific Studies of Reading*, 26(1), 1-20. <https://doi.org/10.1080/10888438.2020.1850733>