JOSEA
THE JOURNAL OF SPECIAL EDUCATION APPRENTICESHIP
ISSN 2167-3454
Vol. 7(1) Feb 2018

The Influence of Student Characteristics on Early Elementary Oral Reading Fluency

Wilhelmina van Dijk
University of Florida

Oral Reading Fluency (ORF) is a widely-used index of reading ability in early elementary grades; however, little information exists on predictive value of student characteristics on ORF scores (Wang, Algozzine, Ma, & Porfeli, 2011). A three-step sequential model was used to analyze the influence of student characteristics on scores (N = 2649) on an end of year ORF measure. Results indicate gender, race, lunch status, and English Language Learner status explained 7% of the variance in scores after controlling for grade and school characteristics (ΔR^2 = .07, $F_{8, 2626}$ = 35.93, p = < .001), and Special Education (SPED) status explained an additional 5% (ΔR^2 = .05 $F_{9, 2625}$ = 59.45, p = < .001). The predictive value of several student characteristics changed depending on SPED status, and this was also a significant moderator on grade level (ΔR^2 = .002 $F_{2, 2623}$ = 4.12, p = .016). The use of these results in subsequent research is discussed.

Keywords: oral reading fluency, mediators, sequential regression

Oral reading fluency (ORF; i.e., reading with automaticity, accuracy, and expression) is highly related to reading comprehension and to reading proficiency in general (Hudson, Lane, & Pullen, 2005). Its validity as an indicator of reading ability was established in the 1980's (Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988). Since then, ORF has been studied more extensively than any other curriculum-based measure (Baker et al., 2008). Some of the more recent foci of this research relate to predictive validity of ORF on high-stakes statewide assessments and other standardized measures of reading ability (e.g., Baker et al., 2008, Yeo, 2010), the use of ORF passages as intervention stimuli (Stevens, Walker, & Vaughn, 2017), and as predictors of reading difficulties (e.g., Burns, Silberglitt, Christ, Gibbons, & Coolong-Chaffin, 2016). However, only a few extensive studies of early elementary age students' ORF abilities exist. In fact, Wang, Algozzine, Ma, and Porfeli (2011) assert that "it is surprising that there are few large-scale, structured assessments that have clearly described what young students' oral reading rates are, how they change over time, and what external variable impact changes in them" (p. 442). This study aims to fill part of this gap and extend knowledge of the use of measures of ORF by looking at the influence on young readers' ORF scores of (a) student

characteristics, (b) the interaction between special education (SPED) status and grade level, and (c) the interaction between SPED status and gender.

Student Characteristics and ORF

In their large-scale study examining the growth of second grade students' ORF, Wang et al. (2011) focused on gender, special education (SPED) status, and school conditions as potential moderators of reading rate. The results from their study indicated that students receiving SPED services scored significantly lower at all three measurement points in the year (i.e., fall, winter, spring). Moreover, their growth across the year was significantly less than students who did not receive SPED services. Additionally, these authors found gender to be a reliable predictor of ORF scores, with girls outperforming boys. Finally, several school characteristics served as valid predictors: the percentage of students receiving free or reduced priced lunch (FRPL), school size by enrollment, the percentage of European American students enrolled, and a high average total reading score.

In a second large-scale study, Wanzek, Otaiba, and Petscher (2013) investigated the difference in reading growth between students receiving SPED services (in particular those receiving services for emotional/behavioral disability and specific learning disabilities) and their general education peers, and between male and female students. Their sample consisted of students in both second and third grade. The outcomes of this study are in line with those of Wang and colleagues (2011) and show students receiving SPED services averaged significantly lower ORF scores and had a lower growth rate. Additionally, the results indicated gender was a moderator for ORF scores, with females scoring higher on average.

Klein and Jimerson (2005) set out to examine potential bias in ORF as a predictive measure for standardized achievement test. The variables in their study were ethnicity, gender, language background, and socio-economic status. Their sample included Hispanic and Caucasian students in grades 1-3. The results of their analysis suggested no bias based on gender or socio-economic status. Language background and ethnicity, however, did serve as biased predictors, favoring English speakers and Caucasian students.

The outcomes of these studies indicated race, gender, lunch status, language status, and SPED status were significant predictors for students' ORF scores, with some contention between race and lunch status. It is unclear, however, to what degree these factors influence ORF scores when school level characteristics and grade level have been taken into account. Additionally, since SPED status is both associated with lower scores and a lower degree of growth, it is plausible SPED status moderates the relationship between grade level and ORF scores. Therefore, this study will focus on the following research questions:

1. What are the additional effects of student characteristics (i.e., gender, lunch status, English Language Learner status, and race) on first, second, and third grade students' end of year Oral Reading Fluency scores after controlling for school and grade level?
2. What is the additional effect of Special Education status on first, second, and third grade students' end of year Oral Reading Fluency scores after

controlling for grade level, school, and student characteristics?

3. In what way does Special Education status influence the effect of gender and grade level on first, second, and third grade students' end of year Oral Reading Fluency scores?

Methods

To answer the research questions, I used de-identified student level data from a large-scale study on the effect of professional development with early-elementary grade teachers from a Northeastern state in the application of a curriculum-based measurement. This study included administration of ORF measures three times per year for three years. To limit the influence of original study effects, I only included the end of year (EOY) ORF scores from the last year of data collection of the original study, after all teachers in the study had received the professional development.

Independent Variables

In this analysis, a distinction was made between two types of individual variables. The first set of variables represented attendance differences, i.e., grade level (1= first grade, 2 = second grade, and 3 = third grade) and school (a total of 13 schools were represented in the data). School characteristics can be influential in ORF scores (Wang et al., 2011). The second set of variables were student level characteristics, i.e., gender (1 = Female, 0 = Male), lunch status (1 = FRPL, 0 = Full price), ELL status (1 = Yes, 0 = No), race (1 = Caucasian, 2 = African-American, 3 = Hispanic, 4 = Asian, 5 = Mixed, 6 = American Indian), and SPED status (1 = Receiving services; 0 = Not receiving services).

Dependent Variable

In the original study, teachers received professional development on the application of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) as a curriculum-based measurement. Teachers assessed their students with DIBELS three times a year for three years. This analysis only included the ORF subtest scores administered in the spring of the last year of the study (EOY), after each teacher had received the professional development. The DIBELS ORF probes contain short passages of connected text for each grade level. During administration, students are asked to read at a comfortable pace for one minute. Any hesitations (i.e., a pause of more than three seconds), mispronunciations, and omissions are considered errors; self-corrections within three seconds are not considered errors. At the end of the administration, the number of words read correct per minute (wcpm) is the student's score on the probe. Wcpm are calculated by subtracting all errors from the total words attempted. Each administration involves three probes and the median wcpm score of these three probes is used as the final score. For grades 1-3, the alternate form reliability coefficient on three forms ranges from r = .96 to r = .98, the slope improvement reliability coefficient ranges from r = .55 to r = .82, and interrater reliability for all three grades is r = .99 (Dewey, Powell-Smith, Good, & Kaminski, 2015).

Design and Data Analysis

Of the original 7988 observations in EOY ORF for 2013, I retained observations of those participants that had a value for all of the independent variables. The final data set contained 2649 participant observations, approximately 33% of the original set. See Table 1 for participant characteristics.

Table 1
Participant characteristics.

Group	*N*	%
Total	2649	100
Gender		
Male	1384	52.2
Female	1265	47.8
SPED Status		
No SPED	2345	88.5
SPED	304	11.5
Grade Level		
1st grade	857	32.4
2nd grade	911	34.4
3rd grade	881	33.3
Lunch Status		
FRPL	1717	64.8
No FRPL	932	35.2
ELL Status		
ELL	345	13
No ELL	2304	87
Race		
Caucasian	947	35.7
African-American	420	15.9
Hispanic	1099	41.5
Asian	90	3.4
Mixed	82	3.1

Note: SPED = Receiving special education services. FRPL = Eligible for free or reduced priced lunch. ELL = English Language Learner.

To answer RQ1 & RQ2, I used a three-block sequential regression, specified as

$$y_i = [b_{grade}\, x_i + b_{school}\, x_i] + a + e_i, \quad (1)$$

$$y_i = [b_{grade}\, x_i + b_{school}\, x_i] + [b_{ELL} x_i + b_{lunch}\, x_{i+} b_{race}\, x_i + b_{gender} x_i] + a + e_i, \quad (2)$$

$$y_i = [b_{grade}\, x_i + b_{school}\, x_i] + [b_{ELL} x_i + b_{Lunch}\, x_{i+} b_{race}\, x_i + b_{gender} x_i] + [b_{SPED} x_i] + a + e_i, \quad (3)$$

where *i* represents an individual, *y* represents observed scores on the dependent variable, *x* represents observed values of the independent variables, *b*'s represent slopes for each independent variable, *a* represents an intercept, and *e* represents the residual term. To answer RQ3, I added interaction terms between SPED status and gender and SPED status and grade level first separately, and then combined to model (3)

$$y_i = b_{grade}\, x_i + b_{school}\, x_i + b_{ELL} x_i + b_{FRPL}\, x_{i+} b_{race}\, x_i + b_{gender} x_i + b_{SPED} x_i + b_{gender:SPED} + a + e_i, \quad (4)$$

$$y_i = b_{grade}\, x_i + b_{school}\, x_i + b_{ELL} x_i + b_{FRPL}\, x_{i+} r_{race}\, x_i + b_{gender} x_i + b_{SPED} x_i + b_{grade:SPED} + a + e_i, \quad (5)$$

$$y_i = b_{grade}\, x_i + b_{school}\, x_i + b_{ELL} x_i + b_{FRPL}\, x_{i+} r_{race}\, x_i + b_{gender} x_i + b_{SPED} x_i + b_{grade:SPED} + b_{gender:SPED} + a + e_i. \quad (6)$$

Within each block, all independent variables were entered simultaneously and all categorical independent variables were represented by effect codes to avoid an artificial reference group on the intercept. All models were fitted with the *car* package (Fox & Weisberg, 2011) in *R* (R Core Team, 2013).

Results

Regression model (1) examined the influence of school and grade level on EOY ORF scores. This model was able to explain 28.6% of the variance in scores (R^2 = .29, $F_{14, 2634}$ = 75.49, p = < .001). The addition of student characteristics (i.e., ELL status, FRPL, race, and gender) in model (2) increased the explained variance to 35.7%. This increase represented a statistically significant change in R^2 (ΔR^2 = .07, $F_{8, 2626}$ = 35.93, p = < .001), indicating that these four characteristics together accounted for 7% of the remaining variance not explained by school level characteristics and grade level. The subsequent addition of SPED status in model (3) increased the total R^2 with an additional 5% (ΔR^2 = .05 $F_{9, 2625}$ = 59.45, p = < .001) to 40.7%. Of all student characteristics, SPED status accounted for the most extreme differences in scores, with students with disabilities scoring on average 29.87 wcpm lower than students without disabilities after controlling for all other variables in the model (b_{contr} = 14.94, t = 14.94, p = < .001).

The addition of the interaction terms between SPED status and grade level, gender and a combination of both in models (4), (5), and (6) were all statistically significant (ΔR^2 = .002, $F_{2, 2623}$ = 4.12, p = .016; ΔR^2 = .001, $F_{1, 2624}$ = 4.93, p = .027; and ΔR^2 = .003, $F_{3, 2622}$ = 4.12, p = .006, respectively). Since the last model resulted in a small but statistically significant addition to the model, with an additional 0.3% of variance explained, the final effects of student characteristics were interpreted only after splitting the file by SPED status and running separate simultaneous models for both groups. All assumptions for the models were met, and outlier analysis separated per SPED status did not show any influential data points for either group.

For students without SPED status, the adjusted average EOY ORF score was 86.8 wcpm, the total variance explained 39% (R^2 = .39, $F_{22, 2322}$ = 67, p = < .001), and all predictors were statistically significant for the model. Students receiving SPED services, however, had an adjusted average of 53.4 wcpm, 26.1% of the variance in scores explained (R^2 = .26, $F_{21, 282}$ = 4.74, p = < .001), with only race and lunch status as significant predictors. With regard to the interactions, male students without SPED status scored on average 1.8 wcpm lower than the adjusted average, with the difference between male and female students being approximately 3.6 wcpm (b = -1.83, t = -2.83, p = .005); the difference between male and female students receiving SPED services, while larger (i.e., 5.4 wcpm), was not statistically significant (b = 2.69, t = 1.18, p = .24). See figure 1 for this interaction.

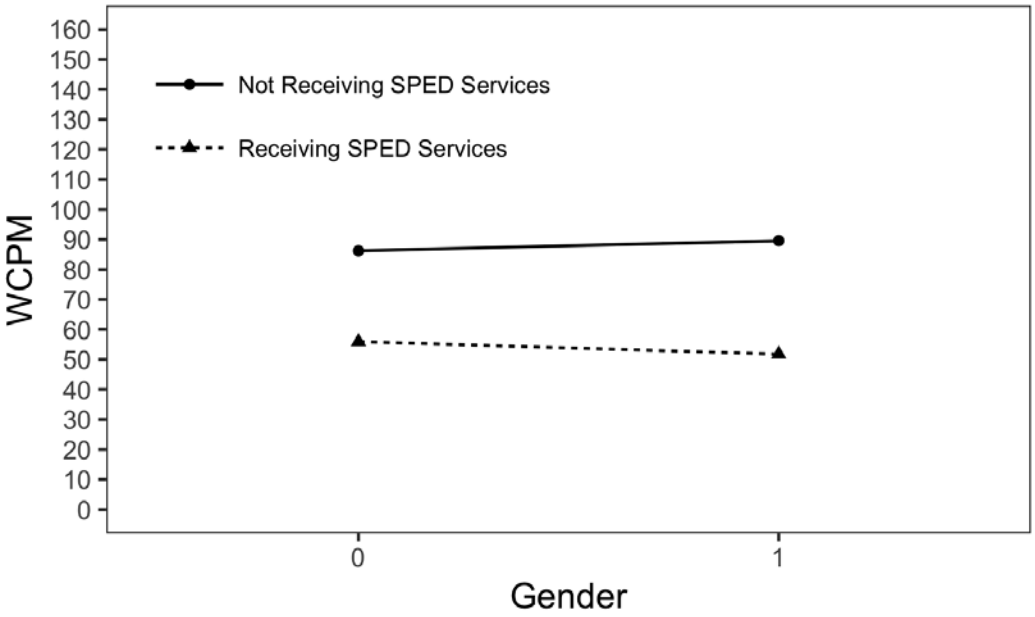

Figure 1. Words read correct per minute (WCPM) as a function of gender and special education status. .

The interaction of SPED status with grade level is shown in Figure 2.

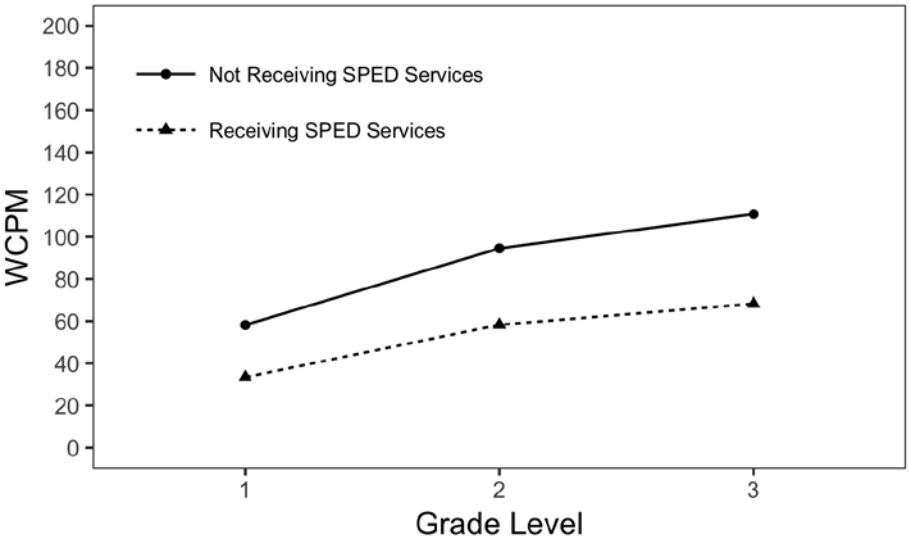

Figure 2. Words read correct per minute (WCPM) as a function of grade level and special education status..

While the general trend of the slopes was similar, the magnitudes were different. Students without disabilities showed an average difference of 34 wpcm between EOY first and second grade, and an average difference of 16 wcpm between EOY second grade and third grade after controlling for all other variables in the model; students with disabilities had an average difference of 26 wcpm from first to second grade, and an average difference of 8 wcpm from second to third grade after controlling for all other variables in the

model (see Table 2 for the disaggregated statistics).

Table 2

Coefficient Estimates of Student Characteristics for Students According to SPED Status.

Group	*SPED*			*No SPED*		
	b	t^a	*p*	*b*	t^b	*p*
Adjusted Grand Mean	53.4	15.06	< .001***	86.8	41.76	< .001***
Gender						
Male	2.69	1.18	.241	-1.84	-3.37	.005**
Female	-2.69	-1.18	.241	1.84	3.37	.005**
Grade Level						
1st grade	-20.08	11.87	< .001***	-28.56	-30.46	< .001***
2nd grade	5.99	2.91	< .001***	6.23	6.80	< .001***
3rd grade	14.1	4.80	< .001***	22.33	24.10	< .001***
Lunch Status						
FRPL	-6.79	-2.40	.017*	-5.66	-7.23	< .001***
No FRPL	6.79	2.40	.017*	5.66	7.23	< .001***
ELL Status						
ELL	-1.06	-0.36	.719	-9.47	-8.82	< .001***
No ELL	1.06	0.36	.719	9.47	8.82	< .001***
Race						
Caucasian	10.81	1.62	.106	17.18	2.16	.031*
African-American	2.16	0.29	.773	-1.69	-0.81	.419
Hispanic	-0.48	-0.08	.939	-8.38	-3.68	< .001***
Asian	-11.06	-.053	.596	-7.62	-3.74	< .001***
Mixed	-1.42	-0.12	.907	9.08	2.72	.006**
American Indian	*NA*	*NA*	*NA*	-8.57	-2.45	.014*

Note. SPED = Receiving special education services. FRPL = Eligible for free or reduced priced lunch. ELL = English Language Learner.

[a]*df* = 282. [b]*df* = 2322.

* *p* < .05, ** *p* < .01, *** *p* < .001.

Discussion

This study examined the influence of student characteristics on EOY ORF scores and the possible moderation of SPED status on these scores. The results of the sequential regression models indicate that SPED status by itself can explain almost as much additional variance in EOY ORF scores as do gender, race, ELL status, and lunch status combined, after taking the variance explained by grade level and school characteristics into account. The outcomes from the interaction analysis of SPED status and grade level indicate that the increase in EOY ORF scores of students with disabilities mimics that of regular education students, with a substantial increase in scores during second grade and a smaller gain during third grade. The magnitude of the increments, however, differed considerably. In general, students with disabilities started with a lower average ORF score in grade 1,

and the increase per grade level was inferior to students without disabilities. This lower growth rate has been presented in previous research, such as Deno, Fuchs, Marston, and Shin (2001). Additionally, even though the outcomes of the current study cannot be taken as a growth model, since each participant provided only one ORF score, the results partly endorse the results from Wanzek et al. (2013) and Wang et al. (2011). Results from these studies showed lower individual growth curves on ORF measures for students receiving special education compared to students receiving general education.

The influence of gender also depended on SPED status. This variable was a significant predictor for students without disabilities, with female students outperforming male students. The difference between male and female students did not have a great practical significance, however, with male students scoring on average 3.8 words per minute less than females. On the other hand, gender was not a significant predictor of ORF scores for students with disabilities. This discrepancy may be due to the difference in sample size in participants with and without disabilities. The small number of students in the SPED group (n = 304) might not have provided that model with sufficient power to detect differences between the gender groups, while the larger sample size in the general education group (n = 2345) did result in adequate power. This could have been the reason for the differences of statistically significance between the models. Therefore, the influence of gender on ORF scores reported by other authors (e.g., Wang et al., 2011; Wanzek et al., 2013) could not be confirmed based on this analysis.

Limitations and Directions for Future Research

This study has several limitations. First, this analysis did not look at the predictive value of ORF scores for the different groups of students on standardized reading measures, which are often used as indicators of students' overall reading competence (Hudson et al., 2005). Thus, it cannot provide information on potential predictive bias of ORF scores based on SPED status. Structural equation models could evaluate the influence of student and school characteristics on ORF scores and their indirect influence on reading outcomes. Since it is important, however, to include all relevant predictors or moderating variable in different statistical models (Kline, 2016), this study may help future research by providing additional evidence on influential characteristics. Additionally, future research could look at the ORF scores for all assessment periods of each year and use more advanced statistical modeling techniques to examine for example growth curves, or the difference in growth between groups through invariance testing. A second limitation was the handling of missing data. The ordinary least squares estimation of regression uses list-wise deletion (Field, 2013), which reduced the number of respondents in our sample considerably. Other statistical techniques, such as multi-level modeling and structural equation modeling can handle missing data better either through pairwise deletion or full information estimators (Enders, 2001). It is possible that the use of these estimators may change the outcomes of the research questions.

Third, this study analyzed secondary data. Little information was available on implementation fidelity of the curriculum-

based measure by the teachers, the professional development teachers received, or instructional decisions made according to the data. Additionally, there was no information on disability category. As such, scores from all students receiving SPED services aggregated in one group. Future research could look at differences between disability categories. Finally, school characteristics were treated as one of the control variables in the models tested. All school characteristics were taken together as one variable. It might be of interest, however, to examine school characteristics in more detail, by looking at enrollment specifics, a school's potential (as defined in Wang et al., 2011), or the nature of literacy instruction. Additionally, modeling the outcomes by school, grade level, and SPED status could identify variation in performance across schools. Studying this variation in depth could establish specific, successful implementation approaches that might help to increase effectiveness across schools through more individualized professional development (Bryk, Gomez, Grunow, & LeMahieu, 2015). Combining successful implementation approaches and effective interventions may hold the key to improving the ORF of students receiving SPED services.

Conclusion

The results from this analysis show SPED status can explain a large amount of the variation in the end of year oral reading fluency scores, even after the influence of grade level, student level, and school characteristics are taken into account. Furthermore, SPED status was a significant moderator for grade level and gender. Subsequently, more sophisticated statistical models aiming to examine the relationships between the different predictors, both at the school and at the student level, should include SPED status as a moderator to obtain more precise estimates of the influence of each predictor.

References

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'Enui, E. J., & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review*, *37*(1), 18-37.

Bryk, A., Gomez, L., Grunow, A., & LeMahieu, P. (2015). Learning to improve: How America's schools can get better at getting better. Cambridge, MA: Harvard Education Press.

Burns, M. K., Silberglitt, B., Christ, T. J., Gibbons, K. A., & Coolong-Chaffin, M. (2016). Using oral reading fluency to evaluate response to intervention and to identify students not making sufficient progress. In: K. Cummings, & Y. Petscher. (eds) *The Fluency Construct* (pp. 123-140). New York, NY: Springer.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurements to establish growth standards for students with learning disabilities. *School Psychology Review*, *30*(4), 507-524.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, *61*(5), 713-740.

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.) Thousand Oaks, CA: Sage Publications.

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd ed.). Thousand Oaks, CA: Sage.

Fuchs, L., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-28.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills (6th ed.).* Eugene, OR: Institute for the Development of Educational Achievement. Available: http://dibels.uoregon.edu/.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher, 58*(8), 702-714. doi:10.1598/RT.58.8.1

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.) New York, NY: Guilford.

Klein, J. R. & Jimerson, S. R. (2005). Examining ethnic, gender, language, and socioeconomic bias in oral reading fluency scores among Caucasian and Hispanic students. *School Psychology Quarterly, 20*(1), 23-50. doi: 10.1521/scpq.20.1.23.64196

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Stevens, E. A., Walker, M. A., & Vaughn, S. (2017). The effects of reading fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: A synthesis of the research from 2001 to 2014. *Journal of Learning Disabilities, 50*(5), 576-590. doi: 0022219416638028.

Wang, C., Algozzine, B., Ma, W., & Porfeli, E. (2011). Oral reading rates of second-grade students. *Journal of Educational Psychology, 103*(2), 442-454. doi:10.1037/a0023029

Wanzek, J., Al Otaiba, S., & Petscher, Y. (2013). Oral reading fluency development for children with emotional disturbance or learning disabilities. *Exceptional Children, 80*(2), 187-204.

Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education*, *31*(6), 412-422.

Author Note:

Wilhelmina van Dijk, School of Special Education, School Psychology, and Early Childhood Studies, University of Florida.

Correspondence to this article should be addressed to Wilhelmina van Dijk, School of Special Education, School Psychology, and Early Childhood Studies, University of Florida, Gainesville, FL 32611. Email: willavandijk@ufl.edu