Set for variability as a critical predictor of word reading : Potential implications for early identification and treatment of dyslexia

Steacy, Laura M., Edwards, Ashley A., Rigobon, Valeria M., Gutiérrez, Nuria, Marencin, Nancy C., Siegelman, Noam, Himelhoch, Alexandra C., Himelhoch, Cristina, Rueckl, Jay and Compton, Donald L.

Set for Variability as a Critical Predictor of Word Reading: Potential Implications for Early Identification and Treatment of Dyslexia

Laura M. Steacy[1], Ashley A. Edwards[1], Valeria M. Rigobon[1], Nuria Gutierrez[1], Nancy C. Marencin[1], Noam Siegelman[2,4], Alexandra Himelhoch[1], Cristina Himelhoch[1], Jay Rueckl[2,3], and Donald L. Compton[1]

[1]Florida Center for Reading Research, Florida State University, Tallahassee, FL USA

[2]Haskins Laboratories, New Haven, CT USA

[3]University of Connecticut, Storrs, CT USA

[4]Departments of Psychology and Cognitive and Brain Sciences, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel

This research was supported in part by Grants P20HD091013 and R21HD108771 to Florida State University by Eunice Kennedy Shriver National Institute of Child Health and Human (NICHD) and by Grant R324B190025 awarded to Florida State University by the Institute of Education Sciences (IES). The content is solely the responsibility of the authors and does not necessarily represent the official views of NICHD or IES.

Correspondence concerning this article should be sent to Laura M. Steacy: lsteacy@fcrr.org.

Abstract

Quasiregular orthographies such as English contain substantial ambiguities between orthography and phonology that force developing readers to acquire flexibility during decoding of unfamiliar words, a skill referred to as a "set for variability" (SfV). The ease with which a child can disambiguate the mismatch between the decoded form of a word and its actual lexical phonological form has been operationalized using the SfV mispronunciation task (e.g., the word *wasp* is pronounced to rhyme with *clasp* [i.e., /wæsp/] and the child must recognize the actual pronunciation of the word to be /wɒsp/). SfV has been shown to be a significant predictor of word reading variance. However, little is known about the relative strength of SfV as a predictor of word reading compared to other well-established predictors or the strength of this relationship in children with dyslexia. To address these questions, we administered the SfV task to a sample of grade 2-5 children (N=489) along with other reading related measures. SfV accounted for 15% unique variance in word reading above and beyond other predictors, whereas phonological awareness (PA) accounted for only 1%. Dominance analysis indicated SfV is the most powerful predictor, demonstrating complete statistical dominance over other variables including PA. Quantile regression revealed SfV is a stronger predictor at lower levels of reading skill, indicating it may be an important predictor in students with dyslexia. Results suggest that SfV is a powerful and potentially highly sensitive predictor of early reading difficulties and, therefore, may be important for early identification and treatment of dyslexia.

Set for Variability as a Critical Predictor of Word Reading:Potential Implications for Early Identification and Treatment of Dyslexia

During the earliest stages of reading development, children are tasked with establishing and storing precise orthographic representations that allow automatic word reading. Precise orthographic representations are "deterministic" (Perfetti, 1991) and impenetrable to factors such as knowledge and expectation (Perfetti, 2017); thus, permitting an orthographic input to sufficientlyand uniquely identify the word to be read (see Castles et al., 2018). This in turn activates phonological, syntactic, morphological, and semantic information to be used by the reader to form faithful representations of text (see Foorman et al., 2016; Kintsch & Rawson, 2005; Snow, 2002). Several studies have reported that relatively few successful exposures to a word are required for the establishment of an orthographic representation in typically developing readers (e.g., Brooks, 1977; Ehri & Saltmarsh, 1995; Reitsma, 1983), with significant variation in the efficiency of item-level orthographic learning as a function of both child- and word-factors (e.g., Duff & Hulme, 2012; Ehri & Saltmarsh, 1995; Nation & Cocksey, 2009; Seidenberg et al., 1984; Steacy & Compton, 2019; Taylor et al., 2011). This evolution from novice to skilled word reading is marked by the transition from deliberate and effortful letter-by-letter decoding to an understanding of, and ability to unconsciously exploit, the complex and probabilistic relationships between orthography and phonology which is the hallmark of skilled word reading and decoding (see Harm & Seidenberg, 2004; Perfetti, 2017; Seidenberg, 2005; Seidenberg & MacDonald, 1999; Steacy et al., 2019a; Treiman et al., 2006).

In establishing complex connections between orthography and phonology needed to simultaneously decode and decipher words (see Gough et al., 1992), developing readers build knowledge of orthographic-to-phonological correspondences that are sensitive to the irregularities of the English language. For instance, Steacy et al. (2019a) have reported that the probability of a developing reader using "context-dependent" vowel pronunciations (e.g., pronouncing the nonword *zead* to rhyme with *head*) is predicted by the child's general reading ability and support for -*ead* pronounced as /ɛd/ within the written English corpus. These results are consistent with computational models of reading, in this case the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996), in which mappings between orthography and phonology are represented in a distributed manner across the reading network and come to represent the probabilistic structure of the corpus rather than simply being rule-based (see Mousikou et al., 2017; Seidenberg, 2005). As developing readers encounter longer (e.g., polysyllabic) words, they are confronted with more varied and complex inconsistencies in spelling-to-sound mappings and are presented with new probabilistic learning challenges such as syllabification, identification of morphological boundaries, stress assignment, and vowel reduction (see Chetail et al., 2015; Mousikou et al., 2017; Perry et al., 2010; Yap & Balota, 2009). As Seidenberg (2017) puts it, "[r]eaders become orthographic experts by absorbing a lot of data, which is one reason why the sheer amount and variety of text that children read is important. For a beginning reader, every word is a unique pattern. Major statistical patterns emerge as the child encounters a large sample of words, and later, finer-grained dependencies such as the fact that syllables can both begin and end with *st,* begin but not end with *tr*, end but not begin with *bs*, and *sb* can only occur across syllables (e.g., *disbar*)" (p. 92). A major pursuit in reading research has been to identify and explain the processes that enable children to efficiently move from novice to skilled word readers (see Adams,1990; Castles et al., 2018; Seidenberg, 2017) along with determining what explains individual differences in this important learning process. We continue this tradition by examining the role of "set for variability" as an important predictor of, and potential mechanism for, word reading development across the distribution of developing readers.

Our intent here is to provide a rationale and empirical support for the important role of a *newly* revived but *older* construct known as “set for variability” (SfV) in facilitating word reading development in developing readers. SfV was first coined by Gibson and Levin (1975) and later resurrected by Venezky (1999). As Venezky puts it, “In learning decoding patterns, children must acquire what Gibson and Levin call a *set for variability*. That is, if one pronunciation does not produce a known word that makes sense for the context, a child has to try a different pronunciation…why some children acquire this set for diversity quickly and others only after great effort remains a mystery” (p. 232). We have spent the last five years exploring SfV as a construct in an attempt to better understand how and why it explains individual differences in word reading and decoding skill in developing readers. To adequately conceptualize our results we first provide a model of word reading development in which SfV plays a prominent role in bridging the gap between the product of decoding (referred to here as the *decoded form*[1]) and the word’s actual phonological representation (referred to here as the *phonological form*); next we present how the concept of SfV has been operationalized into a continuous measure to be used in the study of individual differences in reading; followed by a brief review of the empirical data supporting the relationship between SfV and word reading development. We then present results from a large study of 489 developing readers who were administered the SfV task along with a battery of other reading and reading related measures. Here we try to probe more deeply the unique contribution of SfV in explaining word reading skill, as well as how the relationship between SfV and word reading varies as a function of word reading ability. We use the discussion to explore plausible explanations for the strong relation

[1] In this case, the terms *decoded form* and *spelling pronunciation* can be thought of as synonymous.

between SfV and word reading skill and speculate on the possible role of SfV in the early identification and perhaps treatment of dyslexia.

Orthographic Learning as the Driving Mechanism in Word Learning

The orthographic learning hypothesis (see Castles & Nation, 2006; Nation & Castles, 2017) posits that the transition from novice to skilled word reading involves the continuous addition of fully specified word-specific representations to the orthographic lexicon. Orthographic learning is an item-based acquisition theory that relies heavily on the application of phonological decoding skills to novel printed words via self-teaching (see Share, 1995; 2011), which results in the formation of stable word-specific orthographic representations (see Castles et al., 2018). Item-based acquisition models, such as the orthographic learning hypothesis, acknowledge that at any point in time a child may be reading some words slowly and with great effort, while other words are read automatically (Castles & Nation, 2006; Share, 1995), with item-level variation likely depending on individual differences in factors such as the frequency and richness of reading (Cunningham & Stanovich, 1998), phonological decoding skill (de Jong et al., 2009; Nation & Castles, 2017), and semantic knowledge (Ouellette & Fraser, 2009).

While phonological decoding skill is surely necessary to support orthographic learning, it is not sufficient to guarantee the formation of a particular word-specific representation (see Nation & Castles, 2017; Share, 2008). At the level of the item (i.e., the specific word to be learned) there are word features and child characteristics that either promote or inhibit orthographic learning. For developing readers, attempting to decode an unfamiliar letter string can result in either full or partial decoding (see Castles & Nation, 2006; Elbro et al., 2012; Keenan & Betjemann, 2008; Tunmer & Chapman, 2012; Venezky, 1999). Full decoding occurs when the reader has sufficient decoding skills to sound out the word and the word contains regular (or decodable) relationships between orthography and phonology. Partial decoding, on the other hand, occurs when the reader does not have sufficient decoding skills to sound out the word, or the word is irregular and cannot be pronounced correctly by applying common decoding rules (e.g., *was*, *have*, *come*, *said*, *kind*, *shoe*, *wasp*, *stomach*, *soup*, *iron*, etc.; Wang et al., 2013). During full or partial decoding, the role of the reader is to match the assembled decoded form of the letter string with the stored phonological form of the word (see Share, 2008; Venezky, 1999). Thus, the probability that a child will correctly decode an unfamiliar letter string depends on the decoding knowledge of the reader, the regularity of the orthographic-to-phonological relationships of the word, and the ability of the reader to bridge any differences between the decoded form and the correct phonological form stored in memory. The "distance" between the decoded and actual phonological form of the word affects the probability that a developing reader will successfully make the match between them (see Edwards et al., in press). Further, the availability of top-down support, either through activation of the stored phonological form (e.g., Duff & Hulme, 2012; Wang et al., 2013) or meaning (Ouellette & Fraser, 2009), likely aids a child in determining the exact pronunciation of a novel letter string on the basis of a partial decoding attempt, suggesting lexical support in orthographic learning under conditions of decoding ambiguity (see Wang et al., 2012, 2013). Elbro et al. (2012) offered that the ability to match the decoded form of the letter string with the stored phonological form of the word serves as a bridge between decoding and lexical pronunciations and may be an important second step in the decoding process. As such, orthographic learning is relevant to the learning of all words with differences in the speed of a child acquiring a reliable orthographic representation for a given word being influenced by a combination of the child's decoding ability and the availability of word meaning and phonological form (see Perfetti, 1992, 2007; Perfetti & Hart, 2002; Perfetti, &

Stafura, 2014); the word's regularity, orthographic complexity, and frequency (see Seidenberg et al., 1984; Waters et al., 1984); and most critically for us here the ability of the reader to bridge the gap between the decoded and phonological form.

Operationalization of SfV

The ability of a child to disambiguate the mismatch between the decoded form of a word and its actual lexical pronunciation has been operationalized in the Set for Variability (SfV) mispronunciation task. The SfV task is a purely oral language task (see Tunmer & Chapman, 1998) requiring an individual to disambiguate the mismatch between the decoded form of an irregular word (provided orally by the assessor) and its actual lexical pronunciation (produced by the child). In the task, children are asked to identify the correct pronunciation of spoken English words that are "mispronounced" by the assessor based on regular decoding rules (e.g., /brikfəst/ for /brɛkfəst/; /wæsp/ for /wɒsp/). Elbro et al. (2012) reported that mispronunciations in Danish based on spelling pronunciations (i.e., decoded form; e.g., munk ('monk') pronounced with a standard [u] rather than the correct, conditional [ʌ]) were more strongly related to word reading than mispronunciations based on other substitutions (e.g., "telefonen" ('the telephone') mispronounced "deleponen"); for this reason, mispronunciation correction tasks in English have typically relied on using the decoded form of irregular words. Tunmer and Chapman (2012) found that SfV items presented in isolation made a unique contribution to variance in exception word reading, whereas a ratio that included the number of mispronounced words identified in context minus the number identified in isolation, did not. They interpreted this to mean that it is the ability to identify mispronounced words that is responsible for the relation between SfV and word reading rather than the ability to use context to select the correct word among several candidates. They reported significantly stronger correlations between word reading skill and the

SfV format that used isolated item performance and for this reason the SfV items are typically administered in the isolated format. Finally, Kearns et al. (2016) reported that the mispronunciation correction task was best described as unidimensional using confirmatory factor analysis in a sample of 206 grade 2-4 children.

Evidence Supporting the Importance of SfV During Orthographic Learning

In terms of the strength of the relationship between SfV and word reading, the literature to date suggests that there is a very strong relationship between SfV and word reading skill across a wide range of ages/grades in developing readers. Studies indicate that this ability to go from a decoded form of a word to a correct pronunciation predicts individual differences in general word reading (e.g., Dyson et al., 2017; Edwards et al, in press; Kearns et al., 2016), irregular words (Steacy et al., 2019b; Tunmer & Chapman, 2012), regular words (Elbro et al., 2012), and nonwords (Kearns et al., 2016; Steacy et al., 2019a; Tunmer & Chapman, 2012). Specifically, in a sample of 5-6 year old children, Tunmer and Chapman (2012) reported predictive correlations (i.e., between SfV in Year 1 and reading measures in Year 3) of .51 between SfV and decoding, .54 between SfV and exception word reading, and .59 between SfV and context word reading. Importantly, Tunmer and Chapman also reported that SfV skill in this sample of children completely mediated the relationship between child vocabulary knowledge and word reading skill. Elbro et al. (2012) found that SfV mispronunciation correction was related to regular word reading in 1st grade Dutch children and longitudinally related to both regular and irregular word reading in Danish children. Steacy et al. (2019b) also reported that SfV was a strong predictor of irregular word reading at both the level of the item and the measure with correlations between SfV and word reading above .60. Edwards et al. (in press) reported correlations of .79 and .76 between SfV and untimed word reading and nonword reading, respectively; with Kearns et al. (2016) reporting a correlation of .62 with untimed real wordreading, .60 with timed word reading, and .59 with decoding of nonwords. While slight variations exist among studies in the reported magnitude of the relationship between SfV and word reading, across studies the data suggest a strong and consistent relationship between the two skills, signifying the importance of SfV as a potential mechanism supporting the transition from novice to skilled word readers.

Present Study

Given the emerging evidence supporting a strong relationship between SfV and word reading, the goal of the present study was to explore the strength of SfV as a predictor of untimed word reading compared to other important predictors of early reading, and to explore its relationship with word reading along the word reading skill distribution in developing readers. In particular, we were interested in using dominance and random forest analyses to compete SfV against other important word reading predictors (i.e., phonological awareness, rapid automatized naming, vocabulary, and attention). In this study, we explore the relative contribution of SfV as a predictor of word reading across the entire distribution of reading skill (using quantile regression), with a particular interest in students at the lower end of the distribution (i.e., children with dyslexia). We take a dimensional approach to exploring the SfV–word reading relationship in children with dyslexia by examining variation as a function of word reading skill, which does not require the reading distribution to be divided into distinct groups representing dyslexic and typically developing readers.

Method

Participants

A sample of 489 grade school children (grades 2-6) was recruited from public elementary schools (N = 320) in the Southeastern United States and from four private schools (N = 169) serving students with learning disabilities along the east coast[2]. Within our sample, 51.1% identified as female. The mean age of participants was 9.17 years. 167 students in 2nd grade, 140 3rd grade students, 141 4th grade students, 36 5th grade students and 5 6th grade students were included in the sample. Additionally, demographic data on the sample indicatedthe following breakdown of racial categories: 40% African American, 43% White, 14% Hispanic, 1% Asian, and 2% identifying as Multiracial. 11% of the sample were English Language Learners. Students with intellectual disabilities were excluded from this study. Demographic information for the sample broken out by school is provided in Table 1.

Measures

Set for variability (SfV). Based on the work of Tunmer & Chapman (1998; 2012), set for variability was assessed by participants' ability to determine the correct pronunciation from spoken words that were "mispronounced" based on common decoding rules, as they might be if they were regularized or partially decoded (e.g., /*brikfəst*/ for /*brɛkfəst/*)[3]. The instructions for this task were: "Today we're going to play a word game with Alex. In this game, Alex tries to say a word but he says it the wrong way. You have to figure out what Alex is trying to say. When Alex says a word, you try to guess what he is trying to say…When Alex says a word, you try to guess what he is trying to say." Note that two examples, with feedback, were provided (*breakfast* and *mother*). A total of 75 items wereadministered and coefficient alpha for our sample was .91.

Phonological awareness (PA). The Elision task from the Comprehensive Test of Phonological Processing was used to test phonological awareness by having participants remove

[2] This sample has previously been used to explore child- and word-level predictors of SfV item performance (Edwards et al., in press). Many of these students either had idenitifed dyslexia or poor word reading.
[3] Note that there was only one overlapping item (*island*) between this SfV task and the word reading items on the WJ-III WID measure.

certain phonemes from words (Wagner et al., 2013). Correct deletion of phonemes and subsequent correct pronunciation served as an indicator of high phonological awareness. The authors report test-retest reliability of .93 (Wagner et al., 2013).

Rapid automatized naming (RAN). Rapid automatized naming was measured using the letter naming task from the Comprehensive Test of Phonological Processing (Wagner et al., 2013). Scores were derived from the amount of time, measured in seconds, that it took participants to correctly name a sequence of letters. If a student could not name all of the practice items correctly, even with error correction, the test was discontinued. If a student made more than 4 uncorrected errors on the test, the score was not used. Test-retest reliability is .72 for children of ages 8–17 years according to the test manual (Wagner et al., 2013).

Attention (ATTN). Participants' general level of attention was measured through teacher evaluations on items 1 through 9 on the Strengths and Weaknesses of ADHD-Symptoms and Normal Behavior Scale (SWAN; Swanson et al., 2012). Scores are scaled such that lower scores indicate more inattention behaviors. Ordinal alpha for these items in this sample was .96.

Vocabulary knowledge (VOC). Vocabulary knowledge was measured using the vocabulary portion from the Wechsler Abbreviated Scale of Intelligence (Wechsler, 2011), in which participants had to name visual objects and provide definitions for presented words. Interrater reliability for elementary age children ranges from .92–.94 (McCrimmon & Smith, 2013).

Word identification skill untimed (WID). Students' ability to identify words was measured using the Word Identification component of the Woodcock Reading Mastery Tests-Revised/Normative Update (Woodcock, 1988). This task was not timed and required that each student read aloud words one by one. An item was only marked correct if the accurate pronunciation was given. The authors report a split-half reliability of .93-.96 for the ages assessed here (Woodcock et al., 2001).

Procedures

Informed consent forms were given to all guardians of the sampled students and the student's assent to be a participant of the study was also received. The tasks were administered by trained research assistant, who had achieved 80% procedural fidelity prior to testing participants in a single, 45-minute session. Results from each task were scored by two research assistants and entered twice in order to minimize human error. Additionally, interrater reliability was assessed by evaluating twenty percent of sessions.

Data Analysis

Several analyses were conducted to evaluate the importance of SfV in the prediction of word reading[4]. First, the amount of variance in word reading explained by SfV that cannot be explained by other common predictors was assessed using hierarchical regression. We specially estimated the unique variance explained by SfV above that which can be explained by vocabulary knowledge, RLN, attention, and PA.

Although hierarchical regression allows for the estimation of unique additional variance explained by each predictor, interpreting the importance of individual predictors is difficult when there is a high degree of multi-collinearity among the predictors. Dominance analysis, an extension of multiple regression developed by Budescu (1993), addresses this issue, allowing for the investigation of the relative importance of correlated predictors. Thus, dominance analysis

[4] Given the makeup of the sample in this study, we ran analyses on the entire sample and broken down by school type (public vs. specialized) and by grade. We did not find substantial differences between these subsamples, and therefore include only the full analyses here. Note that there is not a significant interaction between SfV and grade predicting WID, which indicates that there is not a significant difference in correlation between SfV and WID based on grade.

was also employed to evaluate the importance of SfV in the prediction of word reading.

Dominance analysis estimates the R^2 values of all possible combinations of predictors and uses this to determine if a predictor is dominant over another predictor. Azen and Budescu (2003) divide dominance into three types: complete, conditional, and general. Complete dominance requires that the predictor's additional contribution for each model is greater than that of the competitor predictor. Conditional dominance entails that the average contribution for a model size (number of predictors in the model) is greater for all model sizes for the predictor comparedto the competitor. General dominance is achieved if the predictor's average contribution across all models is greater than that of the competitor predictor. These dominance types are nested such that if complete dominance is achieved then conditional is achieved and if conditional is achieved then general is achieved as well. Dominance analysis was used to assess the dominance structure of these predictors of word reading, specifically the relative importance and dominancestatus of SfV compared to the other predictors.

Linear regression-based analyses (which includes dominance analysis) assume a linear (or curvilinear) association between the predictor and outcome, whereas tree-based analyses allow for nonlinear associations by employing a series of cut-points in order to arrive at an estimated value. In general, variables near the top of the tree are considered to be more important. Random forest allows for a direct investigation of the importance of each variable by examining the increase in error when that variable is permuted (replaced with random noise). The larger the increase in error when that variable is permuted, the more important that variable is to the prediction (i.e., there is not another variable that can take over in its absence). This also provides a rank order of importance of variables to the prediction of word reading. Thus, a random forest was conducted to evaluate the relative importance of SfV in the prediction of word reading in a tree-based analysis, allowing for nonlinear associations.

Lastly, given the importance of SfV to the prediction of word reading, an in-depth look at how the strength of this prediction varies as a function of reading ability is warranted. Quantile regression was used to investigate the strength of the association between SfV and word reading across the distribution of word reading ability. As opposed to ordinary least squares regression which conditionalizes at the mean of the outcome variable, quantile regression allows us to condition at various points along the word reading distribution to investigate whether the impact of SfV on the prediction of word reading differs when conditioned at these different levels of word reading ability. Quantile regression was used to explore how the importance of SfV to the prediction of word reading may differ when conditioned at relatively low, medium, and high levels of word reading ability.

Results

All variables were age-residualized to account for the varying ages of participants to ensure that no observed associations are simply due to age. Outlier values greater than 3 standard deviations from the mean were winsorized, which occurred for 7 cases in RLN, 5 in VOC, and 5 in PA. All 489 participants had complete data on all variables used in this analysis. Means, standard deviations, and correlations (for both raw scores and age residualized values) are presented in Table 2. All predictors were significantly correlated with word reading skill with the strongest correlations being PA (r=.61) and SfV (r=.79). It is worth noting that the correlation between SfV and PA was .66 and SfV and vocabulary was .50 (see Dyson et al., 2017; Kearns et al., 2016; Steacy et al., 2019a for further discussions regarding the relationships between SfV, PA, and vocabulary knowledge). Finally, SfV performance correlated significantly with child attention ratings (r=.33). All variables were also z-scored in all analyses to aid in interpretation.

First, hierarchical regression was used to determine the amount of variance in word reading explained by SfV that cannot be accounted for by vocabulary knowledge, RLN, attention, or PA. Results showed that the model with SfV included explained significantly more variance in word reading than the model without SfV [$F(1,483) = 239.49$, $p<.0001$]. The model including SfV ($R^2 = .6669$) explained an additional 16.5% of the variance than the model without SfV ($R^2 = .5017$). The results from the final model are provided in Table 3. When investigating the unique variance explained by PA over and above the other predictors, the model with all predictors ($R^2 = .6669$) only explained an additional 1.19% of the variance in word reading than the model without PA ($R^2 = .6650$) which was also a significant increase [$F(1,483) = 17.29$, $p<.0001$].

A dominance analysis was conducted in R using the *dominance analysis* package (Navarrete & Soares, 2020) to determine the importance of SfV to the prediction of word reading as well as the relative contribution of the other predictors (VOC, RLN, inattention, and PA). Results showed SfV to have complete dominance over all other predictors, having a greater additional contribution than all other predictors in each model. SfV contributed .33 to the R^2 on average, whereas the next highest was PA which contributed .14 on average, with allother predictors contributing less than .1 on average. This finding suggests that SfV completely dominates both PA and the other predictors in the model.

Next, the relative importance of the predictors of word reading was evaluated using a random forest analysis. A random forest was conducted in R using the *randomForest* package (Liaw & Wiener, 2002) with 100 trees. Overall, the model with all predictors (SfV, vocabulary, RLN, inattention, and PA) explained 61.65% of the variance in word reading. Again, SfV was shown to be the most important predictor, showing the greatest increase in mean square error when permuted as well as the greatest increase in node purity resulting from splits based on a given predictor. Node purity is measured by residual sum of squares, a more pure node means that the cases within a node are more similar on the outcome variable compared to a more impure node where cases are more dissimilar (the mean of the node does not represent all cases well). An increase in node purity results from an informative split in which the resulting two nodes provides a better estimate after the split than prior to the split. An increase in node purity means a decrease in residual sum of squares meaning overall the predicted values are closer to the actual values. Increase in mean squared error and node purity for each variable can be found inTable 4.

Lastly, quantile regression was used to explore whether the strength of the association between SfV and word reading varies as a function of word reading ability. A quantile regression was run using the *quantreg* package in R (Koenker, 2009) with SfV performance predicting word reading at 3 quantiles: .25, .50, and .75. Results showed that the magnitude of the association between SfV and word reading differed across quantiles. A test of equality of distinct slopes showed small but significant differences in slope across quantiles [$F(2, 1465) = 3.08$, p=.0465]. Thestrength of the association between SfV and word reading was higher when conditionalized at the lower end of word reading (see Figure 1). When conditioned at the .25 quantile, the correlation between SfV and word reading was .82 whereas the correlation was .72 when conditioned at the .75 quantile of word reading (see Table 5). Interestingly, SfV was the only predictor that produced this pattern. While RAN and PA did not exhibit differences across the distribution, vocabulary and word reading were more associated at the higher end of the word reading distribution. Thus, SfV may be an especially important predictor for predicting dyslexia.

Discussion

The goal of this study was to further examine the relationship between SfV and word reading and understand this relationship in relation to other predictors and across reading skill. We used a multifaceted approach to investigate the relationship between SfV and word reading using traditional hierarchical regression, dominance analysis, random forest analysis, and quantile regression. The findings support a strong relationship between SfV and word reading and extend the current literature by focusing on the strength of SfV as a predictor when competed against other predictors and how the relationship between SfV and word reading changes depending on reading skill. The results suggest that SfV accounts for a significant amount of unique variance in word reading over and above other important early predictors of reading. Comparatively, the additional variance explained by SfV was much larger than that explained by PA. When competed against other predictors of word reading, SfV was the strongest predictor, completely dominating PA, RAN, vocabulary, and attention statistically. Of particular interest from the perspective of identifying dyslexia was the relative contribution of SfV across the word reading skill distribution. Results indicate that SfV and word reading were most strongly related at the lower end of the word reading distribution. These results suggest that SfV may be a particularly good predictor of risk for dyslexia. We interpret these results through the lens of a distributed model of word reading and speculate that the SfV mispronunciation task taps into skills necessary for successful word reading beyond phonological awareness and vocabulary.

The Role of Phonological Cleanup in Set for Variability

Given the magnitude of the SfV-word reading relationship and the finding that it statistically dominates PA as a predictor of word reading variance, a greater understanding of what contributes to this task is warranted. Certainly, the task of disambiguating the decoded form to identify the lexical phonological form of the word places primary demands on the phonological system in what can be thought of as a “cleanup” process. Connectionist models of reading (see Harm & Seidenberg, 1999, 2004; Harm et al., 2003; Rueckl et al., 2019) contain a set of phonological cleanup units that contribute to decoding by cleaning up noisy or incomplete phonological representations resulting from orthographic input.This allows hidden units from the orthographic component of the model to provide the phonological system with somewhat incomplete input, akin to the decoded form that a child might produce that captures the regularities in the mapping from orthography to phonology. Elbro et al. (2012) were the first to link the demands of the SfV task with specific aspects of the phonological system in the network architecture of connectionist models when they stated, “An important feature of this phonological network is that, if fully trained, it can determine the correct pronunciation of a word even if the output of the orthographic network is noisy” (p. 357).

The distributed nature of connectionist models allows phonological cleanup units to be influenced by phonological, orthographic, and semantic components of the model, thus providing various sources of information to bear on the problem of disambiguating a mispronunciation. Variation in the efficiency of the phonological cleanup may not only account for individual differences in SfV performance, but it is also thought to be associated with differences in word reading ability, with important implications for how SfV may be used in the identification and treatment of dyslexia. Connectionist modeling studies have shown that disruptions to the phonological cleanup units degrades a network’s reading performance in a way that mimics the reading behavior of individuals with moderate to severe phonological dyslexia; that is, severely impaired nonword reading and moderately impaired exception word reading (Harm & Seidenberg, 2004; Steacy et al., 2021; see also Harm & Seidenberg, 1999; Rueckl et al., 2019).

If the phonological network is impaired, more work has to be done by the hidden units that mediate the mapping between orthography and phonology. This increased workload causes an overfitting problem (see Harm, McCandliss, & Seidenberg, 2003): because the hidden units "memorize" word forms and form "item-specific" representations, they become relatively insensitive to sublexical orthographic-phonological regularities. Thus, we hypothesize that an inefficient phonological cleanup system has direct and indirect consequences for both SfV and reading, for related reasons. First, phonological cleanup is an intrinsic component of both mispronunciation correction and decoding; thus relatively poor cleanup skills would have a detrimental effect on both SfV and word reading. Second, poor phonological cleanup has a negative effect on the development of the component of the reading system responsible for mapping orthographic to phonological forms. This further impairs word (and especially nonword) reading and limits the potential role of orthographic-phonological knowledge in SfV performance (see below).

Behaviorally, the SfV task can be conceived of as drawing on similar mechanisms associated with the phonological cleanup units in the Triangle Model. The high correlation between SfV and PA speaks to a reliance on phonological processing to successfully complete the SfV task (see Elbro et al., 2012). Our findings provide further support for the phonological nature of SfV but suggest that SfV captures variance beyond typical measures of PA. While some of this may be due to measurement, there is strong evidence from this study that SfV captures more variance than a typical phonological task and that it demonstrates complete statistical dominance over PA as apredictor. The additional variance captured by the SfV task may be attributable to influences from other aspects of the reading system (i.e., semantic and orthographic) on SfV task performance.

Semantic and Orthographic Influences on Set for Variability Performance

There is evidence to suggest that semantic knowledge aids in orthographic learning and that contextual information can support orthographic learning through self-teaching (Cunningham et al., 2002; Nation et al., 2007). Share (1995) suggested that the semantic support provided by contextual information and children's ability to use the context to determine exact word pronunciations from partial decoding plays an important role in self-teaching. Further, the important role of semantic knowledge and familiarity in item-specific word recognition has been documented for both mono and polysyllabic words (e.g., Kearns et al., 2016; Taylor et al., 2011; Wanget al., 2013). The correlation between vocabulary and SfV (r=.50) found in the present study suggests that there may be a semantic component to the SfV task, with students with larger vocabularies more likely to be successful with the SfV task. The importance of semantics as a word level predictor in the SfV task is also supported by Edwards et al.'s (in press) finding that there is a significant effect of concreteness on SfV performance. There is certainly some semantic influence in the SfV task and Kearns et al. (2016) characterized the SfV task as measuring "a process that allows readers to take the output of phonological recoding assembled using phonological awareness skills and test it against entries in the phonological lexicon using lexical and sublexical semantic knowledge" (p. 457). Tunmer and Chapman (2012) found that SfV mediates the relationship between vocabulary and word reading at the construct level. This has not, however, been explored at the item-level.

Steacy et al. (2019b) offered that there may be an orthographic component to SfV skill, and further speculated (Steacy et al., 2019a) that learning to read may affect how children approach the SfV task through two related processes. The first might be that as children decode new letter strings, they store the incomplete phonological form that is associated with the lexical form, and this incomplete form is available during the SfV task (for a detailed discussion, see Elbro & de Jong, 2017). The other is that as children become better readers and spellers, they may actively use a phonology-to-orthography pathway to use spelling to disambiguate the mispronunciation. That is, when a child is presented with a mispronounced word in the SfV task, they may use phonology-to-orthography associations to 'translate' the mispronounced spoken word into an orthographic form, a process we refer to as orthographic facilitation, from which they can access the correct phonological form associated with the orthographic form stored in memory. Consistent with this view, Edwards et al. (in press) reported that children with better decoding skills are likely using their knowledge of the varied connections between phonology and orthography to aid in SfV item performance whereas children with poor decoding skills may be relying on phonology and semantic knowledge only. These results suggest that orthography is activated (whether consciously or unconsciously) during the task for those with better decoding skill, thus helping to disambiguate the decoded form of a word to the true phonological representation stored in the lexicon. Our conception of orthographic facilitation is based on the orthographic skeleton hypothesis (Wegener et al., 2018), which suggests that a pronunciation of an unknown irregular word generates a more regular spelling and, in our case, facilitates the target word in the SfV task.

Set for Variability Capturing the Distributed Network of the Reading System

We speculate that the SfV task is capturing more distributed knowledge within the reading network beyond just PA skill. From the perspective of a distributed model of word reading, the SfV task appears to be tapping into all three aspects of the reading network that connects orthography, phonology, and semantics. We hypothesize that SfV may go beyond being the bridge between decoding and lexical pronunciations to capture elements of the distributed reading system. This hypothesis is consistent with Edwards et al.'s (in press) demonstration that children’s knowledge of orthographic-to-phonological relationships affect SfV performance beyond phonological and semantic influences, and with the fact that SfV is both an item-specific and general skill (see Steacy et al., 2019b). Thus, success on some of the SfV items may be the result of orthographic learning and the relationship between SfV and word reading is likely bidirectional. Strong PA skills are required for SfV to impact word reading as a metalinguistic skill while strong word reading skills are needed for orthographic knowledge to impact item-specific SfV performance. We speculate that SfV is a better predictor of reading skill than PA alone because the task is more aligned with what students are expected to do during orthographic learning.

Unresolved Questions Surrounding the SfV-Word Reading Relationship in Dyslexia

The findings from Steacy et al. (2019b) are particularly provocative as we work to better understand variability on the SfV task and how it relates to word reading skill. Steacy et al. asked children in grades 2-5 to read the 40 irregular words that form the basis of the SfV task (e.g., *treasure, spinach, deaf, kind, island, piano, prove, lizard*, etc.). In the study, Steacy et al. modeled the relationship between set for variability (both as item-specific and general predictors) and item-level irregular word reading by decomposing irregular word reading variance into child-, word-, and child-by-word (i.e., item-specific SfV) components. Specifically, models contained relevant characteristics of the child (e.g., SfV, PA, RAN, vocabulary), the word (e.g., frequency, number of letters, concreteness, relative transparency), and the child-by-word (e.g., item level SfV performance) predictors.

Model results presented in Steacy et al. (2019b) further indicated that both item-specific and general child level performance on the SfV task were strong and unique predictors of item-specific irregular word reading. This led us to speculate at that time that the ability to complete the SfV mispronunciation task is both an important general metalinguistic skill related to students' ability to successfully arrive at the correct pronunciation of a word and an item-specific skill for word reading skill (i.e., orthographic learning). We envision the metalinguistic portion of the variance to be related to overall orthographic, phonological, and semantic competency; akin to distributed representations within the Triangle Model as described above. It is the unique variance associated with item-specific performance that we find particularly intriguing. As mentioned, Elbro and de Jong (Elbro et al., 2012; Elbro & de Jong, 2017) have argued that the formation of spelling pronunciations (i.e., what we refer to as the *decoded form*) is an intermediate developmental step between letter-sound decoding and 'sight word' reading which is an accurate way of looking at what is usually called "orthographic" whole word learning. Our results are certainly consistent with this view. In addition, Elbro and de Jong (2017) advocate for the unique importance of stored spelling pronunciations in developing readers by suggesting that spelling pronunciations are stored, and are available for use, in the phonological lexicon. The storing of the decoded form (i.e., spelling pronunciation) in the phonological lexicon has the potential to explain why SfV completely dominated PA as a predictor of early word reading, the increased relationship between SfV and word reading at the lower end of the word reading distribution (i.e., children with dyslexia), and perhaps the limited effects thus far reported in the few studies attempting to train SfV in developing readers as a means to increase reading ability (e.g., Dyson et al., 2017, Zipke, 2016).

The storing of spelling pronunciations in the lexicon as a byproduct of phonological decoding allows for a possible explanation for the advantage SfV has over PA in predicting word reading. Elbro and de Jong (2017) elaborate on a mechanism that favors spelling pronunciations of a word over other mispronunciations when they state, “knowledge of spelling pronunciations may become so entrenched in the mental lexicon that the activation of a spelling pronunciation is just as efficient as (or even more efficient than) activation of a standard phonological representation. Word recognition of the new variant pronunciation becomes automatic with no need for conscious awareness” (p. 183). This would suggest that part of the unique relationship between SfV and word reading skill is a by-product of prior decoding of the SfV target word. This explanation of SfV allows item-specific variance at the word level, due to the process of phonological decoding, to contribute to the absence of a unique contribution of PA in the presence of SfV in the Steacy et al. (2019b) models. In addition, activations of stored spelling pronunciations associated with the SfV items likely activate the spelling pattern (i.e., the orthographic code) which in turn activates the stored phonological form. Again, a process that favors those who are better decoders (see Edwards et al., in press for details), thus driving the strong relationship between SfV and word reading.

Given this explanation for SfV, why would the relationship between SfV and word reading be stronger at the lower end of the word reading distribution (i.e., children with dyslexia)? This seems rather paradoxical; on the surface our previous explanation would seem to suggest that with better decoding and word reading skill one would expect increased SfV performance, and with it a higher SfV-word reading correlation. While it is undoubtedly true that better readers will outperform poor readers on the SfV task this does not mean we should necessarily expect a higher SfV-word reading correlation in typical readers. On the contrary, in struggling readers the SfV task may represent a true measure of phonological cleanup that is less contaminated by the presence of stored spelling pronunciations and feedback from stored orthographic spelling patterns. For children with dyslexia, the SfV task stresses the phonological system without the availability of other types of representations (spellings and spelling pronunciations) and therefore represents a “purer” index of phonological processing as it relates to word reading. In typical readers, the availability of spelling pronunciations and spellings stored in the lexicon (phonological and orthographic, respectively) may diminish the phonological demands of the SfV task which in turn reduces the SfV-word reading correlation as word readingskill increases. Thus, SfV appears to hold great promise as a phonological processing task able toearly identify children who are at risk for developing dyslexia. Specific studies are now needed to directly test this hypothesis.

Finally, several recent training studies have been conducted to test the efficacy of teaching SfV with the intent of improving general word reading skill (Dyson et al., 2017; Zipke, 2016). The idea here is that training SfV will increase children’s ability to be flexible with the output of phonological decoding, increase the probability that the child will arrive at the phonological form of the word, and result in better word reading skill. Both studies provided data challenging this mechanism and instead results suggest that training SfV provides item-specific spelling pronunciations that do not transfer to general word reading improvement. In the Zipke study, SfV was considered a strategy and taught by having children systematically alter the letter sounds when sounding words out until they produced a real word, at which point they were taught to confirm the word meaning with the use of context clues. Beginning readers ($N = 15$) in first and second grade were instructed in the SfV strategy during a series of five one-on-one lessons of 20 to 25 minutes each, while a control group ($N = 15$) was encouraged to decode passages without the SfV instruction. It was hypothesized that this brief training would result in readers who were more persistent and more successful in attempting to read words with irregular spellings. Participants who received the experimental training did make more attempts to read untrained words with irregular spellings; however, there was no effect of training on students' ability to determine the correct pronunciation of mispronounced spoken words (i.e., SfV task) or on their ability to read untrained exception words. This finding suggests that beginning readers were able to learn the strategy but that it did not transfer to the reading of words. We submit that the results of this short-term training study may have resulted in the formation, and potential storage, of item-specific spelling pronunciations that do not produce improved reading of non-trained items.

In a longer training study, Dyson et al. (2017) randomly assigned 84 children, ages 5–7 years, to a mispronunciation correction (i.e., SfV) intervention or control group. Children in the intervention group participated in a 4-week program in which they were taught to correct mispronunciations of spoken words as well as being taught the meanings of those words. Children in the control group received no additional teaching. The intervention group made sizeable gains in their ability to correct mispronunciations on trained words and significant but much smaller gains on mispronunciations of untrained words. Item-level analysis of the effects of mispronunciation training on trained and untrained word reading indicated that general skill on the SfV task significantly predicted both trained and untrained word reading at posttest, but the effects in the intervention group were only present for the trained words. Results suggest that over a longer period of training children may be able to apply this oral language strategy to untrained items, but with only limited success. However, the transfer of this training to reading untaught words was not supported, aligning with the short-term training result reported by Zipke. Results support a mechanism between SfV and word reading development that operates at the item level by supporting orthographic learning through the formation, and possibly storage, of a spelling pronunciation of the word during phonological decoding that aids in the acquisition of a fully specified orthographic representation that is linked to the phonological form.

Conclusions

Our purpose here was to probe the relationship between SfV and word reading in developing readers. Specifically, we were interested in investigating the strength of SfV as a unique predictor of word reading skill and exploring the strength of the relationship along the word reading distribution in developing readers. We report strong age-corrected correlations between SfV and word reading (r=.78) in a sample of 489 children in grades 2-5. As a predictor of word reading variance, SfV was found to demonstrate statistical dominance over other important word reading related measures, most notably PA skill, included in the model. In addition, the SfV-word reading relationship was found to be strongest at the lower end of the distribution. We conceptualize our results within an item-based acquisition model of early word reading (i.e., orthographic word learning) in which phonological decoding results in stored spelling pronunciations (i.e., decoded forms) that are linked to both the phonological form and spelling ofa word. As described by Elbro and de Jong (2017), these stored spelling pronunciations are available for, and facilitate, both word reading and spelling in developing readers. We speculatedthat the SfV task in the bottom end of the word reading distribution may represent a pure measure of phonological cleanup that is less contaminated by the presence of stored spelling pronunciations and feedback from stored spelling representations. Finally, it appears that SfV as a strategy can be taught, but with little effect on untrained items, limiting the generalizability of training. This leads us to conclude that SfV may be a potentially powerful early predictor of risk for dyslexia but may have limited utility to improve the phonological cleanup function associatedwith word reading in children with dyslexia. However, we do see important uses of SfV to support acquisition of fully specified item-specific representations during orthographic learning in all beginning readers, and most specifically children with dyslexia. Given the potential importance of these general conclusions regarding SfV and its relationship to word reading development we suggest the need for further study to address the tenability of our conclusions.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA:The MIT Press

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictorsin multiple regression. *Psychological Methods*, *8*(2), 129.

Brooks, L. (1977). Visual pattern in fluent word identification. In A. S. Reber & D. L. Scarborough (Eds.), Towards a psychology of reading (pp. 95–110). Hillsdale, NJ:Erlbaum.

Budescu, D. V. (1993). Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression. *Psychological bulletin*, *114*(3), 542.

Castles, A. & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Challenges and controversies about word recognition and reading.* London: Psychology Press.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition fromnovice to expert. *Psychological Science in the Public Interest*, *19*(1), 5-51.

Chetail, F., Balota, D., Treiman, R., & Content, A. (2015). What can megastudies tell us aboutthe orthographic structure of English words? *Quarterly Journal of Experimental Psychology*, *68*(8), 1519-1540.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology, 82,* 185–199.

de Jong, P. F., Bitter, D. J., Van Setten, M., & Marinus, E. (2009). Does phonological recoding occur during silent reading, and is it necessary for orthographic learning? *Journal of Experimental Child Psychology*, *104*(3), 267-282.

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge inlearning to read words. *Scientific Studies of Reading, 16,* 504–525.

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction andword meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, *21*(5), 392-407.

Edwards, A., Steacy, L. M., Seigelman, N., Rigobon, V. M., Kearns, V. M., Rueckl, J. R., & Compton, D. L. (in press). Unpacking the unique relationship between set for variability and word reading development: Examining word- and child-level predictors of performance. *Journal of Educational Psychology.*

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal*, *7*(3),295–326.

Elbro, C. & de Jong, P. F. (2017). Orthographic learning is verbal learning. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (169-189). Amsterdam, The Netherlands: John Benjamins.

Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A. M. (2012). From spelling pronunciation tolexical access: A second step in word decoding? *Scientific Studies of Reading*, *16*(4), 341-359.

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner,R., & Wissel, S. (2016). Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016-4008). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://whatworks.ed.gov.

Gibson, E. J., & Levin, H. (1975). The psychology of reading. Cambridge: MA: The MIT press.

Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), Reading acquisition (pp. 35–48). Hillsdale,NJ: Erlbaum.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia:insights from connectionist models. *Psychological Review*, *106*(3), 491.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*(2), 155-182.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading:Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review*, *111*(3), 662–720.

Kearns, D. M., Rogers, H. J., Koriakin, T., & Al Ghanem, R. (2016). Semantic and phonologicalability to adjust recoding: A unique correlate of word reading skill?. *Scientific Studies of Reading*, *20*(6), 455-470.

Keenan, J. M., & Betjemann, R. S. (2008). Comprehension of single words: The role of semantics in word identification and reading disability. In E. Grigorenko (Ed.), Single-word reading: Behavioral and biological perspectives (pp. 191–209). Mahwah, NJ:

Erlbaum Publishers.

Kintsch W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.),The science of reading: A handbook (pp. 209–226). Oxford, UK: Blackwell.

McCrimmon, A. W., & Smith, A. D. (2013). Review of Wechsler abbreviated scale of intelligence, second edition (WASI-II). *Journal of Psychoeducational Assessment, 31*(3), 337–341.

Mousikou, P., Sadat, J., Lucas, R., & Rastle, K. (2017). Moving beyond the monosyllable in models of skilled reading: Mega-study of disyllabic nonword reading. *Journal of Memory and Language*, *93*, 169-192.

Nation, K., Angells, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology, 96*(1), 71-84.

Nation, K., & Castles, A. (2017). Putting the learning in orthographic learning. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 147–168). Amsterdam, The Netherlands: John Benjamins.

Nation, K., & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. Journal of Experimental Child Psychology, 103,296–308.

Navarrete, C. B., & Soares, F. C. (2020). Dominanceanalysis: dominance analysis. *R package version*, *2*(0).

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22.

Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology, 104,* 239–251.

Perfetti, C. A. (1991). Representations and awareness in the acquisition of reading competence. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and itsimplications* (pp. 33-44). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C.Ehri, & R. Treiman (Eds.), Reading acquisition (pp.145–174). Hillsdale, NJ: Erlbaum.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, *11*(4), 357-383.

Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven*, (pp. 51-67). Amsterdam, the Netherlands: John Benjamins.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, *11*, 67-86.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22-37.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive psychology*, *61*(2), 106-151.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child*

Psychology, *36*(2), 321-339.

Rueckl, J. G., Zevin, J. D., & Wolf, H. (2019). Using computational techniques to model and better understand developmental word- reading disorders (i.e., dyslexia). In J. A. Washington, D. L. Compton, & P. McCardle, (Eds.), *Dyslexia: Revisiting Etiology, Diagnosis, Treatment, and Policy* (pp. 57-70). Baltimore, MD: Brookes Publishing Co.

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science*, *14*(5), 238-242.

Seidenberg, M. (2017). *Language at the Speed of Sight: How we Read, Why so Many Can't, andwhat can be done about it*. Basic Books.

Seidenberg, M., & MacDonald, M. (1999). A probabilistic constraints approach to languageacquisition and processing. Cognitive Science, 23, 569–588.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of wordrecognition and naming. *Psychological Review*, *96*(4), 523.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When doesirregular spelling or pronunciation influence word recognition? *Journal of VerbalLearning and Verbal Behavior, 23,* 383–404.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. *Advances in Child Development and Behavior, 36,* 31-82.

Share, D. L. (2011). On the role of phonology in reading acquisition: The self-teaching hypothesis. In S. A. Brady, D. Braze, & C. A. Fowler (Eds.), *Explaining individual* *differences in reading: Theory and evidence* (pp. 45–68), Psychology Press.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation.

Steacy, L. M., & Compton, D. L. (2019). Examining the role of imageability and regularity in word reading accuracy and learning efficiency among first and second graders at risk for reading disabilities. *Journal of experimental child psychology*, *178*, 226-250.

Steacy, L. M., Compton, D. L., Petscher, Y., Elliott, J. D., Smith, K., Rueckl, J. G., … Pugh, K. R. (2019a). Development and prediction of context-dependent vowel pronunciation in elementary readers. Scientific Studies of Reading, 23(1), 49–63.

Steacy, L. M., Petscher, Y., Rueckl, J. R., Edwards, A., & Compton, D. L. (2021). Modeling and visualizing the co-development of word and nonword reading in children from first through fourth grade: Informing developmental trajectories of children with dyslexia. *Child Development.* https://doi.org/10.1111/cdev.13468

Steacy, L. M., Wade-Woolley, L., Rueckl, J. G., Pugh, K. R., Elliott, J. D., & Compton, D. L. (2019b). The role of set for variability in irregular word reading: Word and child predictors in typically developing readers and students at-risk for reading disabilities. *Scientific Studies of Reading*, *23*(6), 523-532.

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., ... Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN rating scales. *The International Journalof Educational and Psychological Assessment, 10*(1), 51–70.

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 37, 60–76. http://dx.doi.org/10.1037/a0020126

Treiman, R., Kessler, B., Zevin, J. D., Bick, S., & Davis, M. (2006). Influence of consonantal context on the reading of vowels: Evidence from children. *Journal of Experimental ChildPsychology*, *93*(1), 1-24.

Tunmer, W. E., & Chapman, J. W. (1998). Language prediction skill, phonological recoding ability and beginning reading. In C. Hulme & R. M. Joshi (Eds.), Reading and spelling:Development and disorder (pp. 33–67). Hillsdale, NJ: Erlbaum.

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studiesof Reading, 16,* 122–140.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of AmericanEnglish Orthography*. New York, NY: Guilford Press.

Wagner, R. K., Torgesen, J. K., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive Test ofPhonological Processing 2*. Austin, TX: Pro-Ed.

Wang, H., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning ofregular and irregular words. *Scientific Studies of Reading, 17*, 369–384. http://dx.doi.org/10.1080/10888438.2012.749879

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition, 12*, 293-305.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II).* San Antonio, TX: NCS Pearson.

Wegener, S., Wang, H. C., de Lissa, P., Robidoux, S., Nation, K., & Castles, A. (2018). Children reading spoken words: Interactions between vocabulary and orthographic expectancy. Developmental Science, 21, e12577. doi: 10.1111/desc.12577

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III tests of achievement.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. Journal of *Memory and Language, 60,* 502–529.Zipke, M. (2016). The importance of flexibility of pronunciation in learning to decode: Atraining study in set for variability. *First Language*, *36*(1), 71-86.

Table 1

Demographic breakdown by school type

	LD schools	Public schools
N	169	320
Gender	42.6% female	55.6% female
Race/Ethnicity		
African-American	6.5%	56.6%
Hispanic	1.8%	20.9%
White	88.2%	18.4%
Asian	0%	1.9%
Multiracial	2.4%	1.9%
American Indian/Alaskan Native	0%	0.3%
Native Hawaiian/Pacific Islander	0%	0.6%
ELL	1.2%	16.6%
Age	9.8 (1.1)	8.8 (1.0)
CTOPP Elision SS	8.7 (2.7)	8.6 (3.0)
WJ Letter Word Identification SS	94.4 (14.6)	104.4 (11.4)
WASI Vocabulary SS	10.9 (2.9)	9.7 (3.2)

Note. Standard deviations are in the parentheses.

Table 2

Means, standard deviations, and correlations

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. Age	9.17	1.13							
2. PA	22.17 (-0.02)	6.72 (6.58)	.14**		.34**	-.22**	.45**	.65**	.61**
3. Attention	37.57 (0.08)	11.45 (11.35)	-.13**	.31**		-.19**	.32**	.38**	.45**
4. RLN	20.63 (-0.05)	5.65 (5.10)	-.22**	-.24**	-.15**		-.16**	-.31**	-.41**
5. VOC	23.86 (-0.02)	6.33 (5.87)	.32**	.48**	.26**	-.22**		.46**	.40**
6. SfV	26.74 (-0.11)	11.17 (10.77)	.26**	.66**	.33**	-.35**	.50**		.78**
7. WID	45.97 (-0.02)	8.20 (7.82)	.30**	.61**	.38**	-.43**	.46**	.79**	

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in parentheses represent the age residualizedvalues, values outside of the parentheses represent raw score values. Correlations above the diagonal represent correlations between age residualized values, whereas those below the diagonal represent those between raw scores. PA = phonological awareness; RLN =rapid letter naming; VOC = vocabulary; SfV = set for variability; WID = word identification.
* indicates $p< .05$; ** indicates $p<.01$.

Table 3

Results of regression analysis predicting word identification

Variable	Estimate	*SE*	*t*-value
Intercept	0.00	0.03	0.00
Vocabulary	0.00	0.03	-0.07
RAN	-0.16	0.03	-5.91***
Attention	0.15	0.03	5.23***
PA	0.15	0.04	4.16***
SfV	0.57	0.04	15.48***

Note. Total R^2 = 0.6669; Unique SfV R^2 = 0.1652; Unique PA R^2 = 0.0119.

Table 4

Variable importance in random forest

Variable	Percent increase in mean squared error	Increase in node purity
Set for variability	23.54	161.07
Attention	8.39	75.67
Rapid letter naming	7.99	65.47
Phonological awareness	7.87	101.35
Vocabulary	3.47	61.71

Note. Increase in node purity refers to the decrease in residual sum of squares from splitting on the variable averaged over all the trees. A split with a large increase in node purity is more informative.

Table 5

Coefficients of quantile regressions

coefficient	.25 quantile	.50 quantile	.75 quantile
(Intercept)	-0.40 [-0.47, -0.34]	0.02 [-0.04, 0.08]	0.42 [0.35, 0.46]
Set for Variability	0.82 [0.77, 0.88]	0.75 [0.71, 0.81]	0.72 [0.66, 0.80]

Note. Intercept and slope (SfV) coefficients are presented for quantile regressions at the .25, .50, and .75 quantiles. All variables were z-scored so the slope is analogous to a correlation coefficient. Values inside the bracket represent the lower and upper bound of the confidence interval.

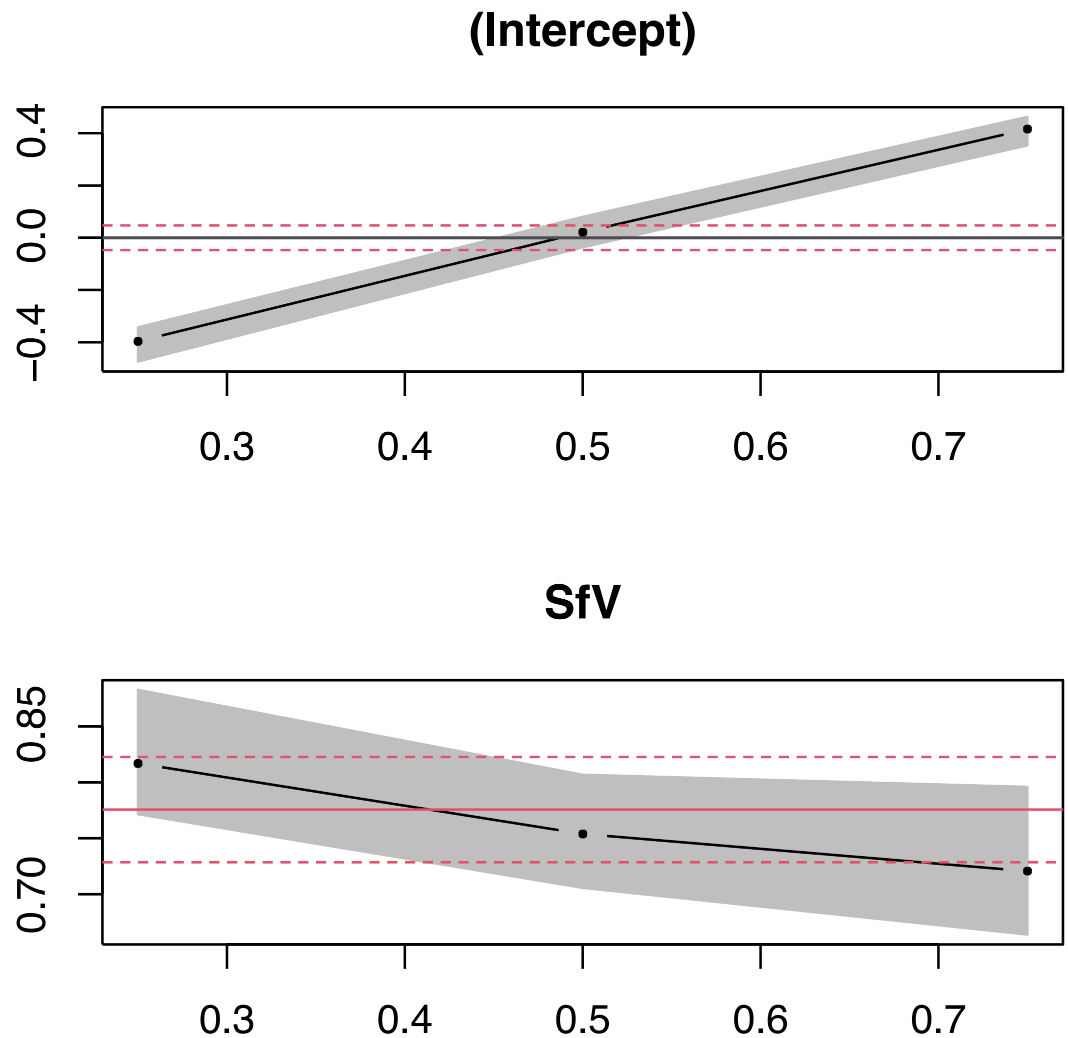

Figure 1. Coefficient values at each of the 3 quantiles (.25, .50, .75) are represented by a black dot, with the surrounding grey area representing the confidence interval. The solid red line represents the coefficient that would be observed in OLS regression when conditioned at the mean, with the dashed red lines represented the upper and lower bounds of the confidence interval of that estimate.

Note. All variables were z-scored so the slope is analogous to a correlation coefficient. SfV = set for variability.