Journal of Educational Psychology

Unpacking the Unique Relationship Between Set for Variability and Word Reading Development: Examining Word- and Child-Level Predictors of Performance

Ashley A. Edwards, Laura M. Steacy, Noam Seigelman, Valeria M. Rigobon, Devin M. Kearns, Jay G. Rueckl, and Donald L. Compton

CITATION

Edwards, A. A., Steacy, L. M., Seigelman, N., Rigobon, V. M., Kearns, D. M., Rueckl, J. G., & Compton, D. L. (2021, October 7). Unpacking the Unique Relationship Between Set for Variability and Word Reading Development: Examining Word- and Child-Level Predictors of Performance. *Journal of Educational Psychology*. Advance online publication. http://dx.doi.org/10.1037/edu0000696

Journal of Educational Psychology

© 2021 American Psychological Association
ISSN: 0022-0663

https://doi.org/10.1037/edu0000696

Unpacking the Unique Relationship Between Set for Variability and Word Reading Development: Examining Word- and Child-Level Predictors of Performance

Ashley A. Edwards[1], Laura M. Steacy[1], Noam Seigelman[2], Valeria M. Rigobon[1], Devin M. Kearns[3], Jay G. Rueckl[2, 4], and Donald L. Compton[1]
[1] Florida Center for Reading Research, Florida State University
[2] Haskins Laboratories, New Haven, Connecticut, United States
[3] Department of Educational Psychology, University of Connecticut
[4] Department of Psychological Sciences, University of Connecticut

Set for variability (SfV) is an oral language task that requires an individual to disambiguate the mismatch between the decoded form of an irregular word and its actual lexical pronunciation. For example, in the task, the word *wasp* is pronounced to rhyme with *clasp* (i.e. /wæsp/), and the individual must recognize the actual pronunciation of the word to be /wɒsp/. SfV has been shown to be a significant predictor of both item-specific and general word reading variance above and beyond that associated with phonemic awareness skill, letter-sound knowledge, and vocabulary skill. However, very little is known about the child characteristics and word features that affect SfV item performance. In this study, we explored whether word features and child characteristics that involve phonology only are adequate to explain item-level variance in SfV performance or whether including predictors that involve the connection between phonology and orthography explains additional variance. To accomplish this, we administered the SfV task (N = 75 items) to a sample of grade 2–5 children (N = 489), along with a battery of reading, reading related, and language measures. Results suggest that variance in SfV performance is uniquely accounted for by measures tapping phonological skill along with those capturing knowledge of phonology to orthography associations, but more so in children with better decoding skill. Additionally, word reading skill was found to moderate the influence of other predictors suggesting that how the task is approached may be impacted by word reading and decoding ability.

Educational Impact and Implications Statement

Set for variability (SfV) is a powerful predictor of word recognition skill in developing readers. The measure taps children's ability to go from the decoded form of a word (e.g., /wʌz/for *was*) to the correct form (e.g., /wɒz/ for *was*), which is considered an important second step in word decoding. In the current study, we worked to determine what factors lead to variability in children's ability to perform the task. We found that performance on the SfV task was highly correlated with children's phonemic awareness skill and also related to their reading and decoding skill. This suggests that children with advanced reading and decoding skill may be using both phonological and spelling skills to go from the decoded form of a word to the correct pronunciation. The findings suggest that further studies evaluating the causal influence of SfV on reading development are warranted.

Keywords: set for variability, word reading, development

Ashley A. Edwards https://orcid.org/0000-0002-5268-4235
Laura M. Steacy https://orcid.org/0000-0003-1603-2990
Noam Seigelman https://orcid.org/0000-0002-5005-4838
Valeria M. Rigobon https://orcid.org/0000-0002-7089-1764
Jay G. Rueckl https://orcid.org/0000-0003-3972-4324
Donald L. Compton https://orcid.org/0000-0003-3855-3746

This research was supported in part by Grant P20HD091013 from the Eunice Kennedy Shriver National Institute of Child Health and Human (National Institute of Child Health and Human Development). Statements do not reflect the position or policy of these agencies, and no official endorsement by them should be inferred.

Correspondence concerning this article should be addressed to Ashley A. Edwards, Florida Center for Reading Research, Florida State University, 1107 West Call Street, Tallahassee, FL 32304, United States. Email: aedwards@psy.fsu.edu

An essential development in learning to read is the acquisition of automatic word reading skills that are impenetrable to factors such as knowledge and expectation (Ehri, 2005; Perfetti, 1992; Stanovich, 1991). 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 that permit an orthographic input to sufficiently and uniquely identify the word to be read (see Castles et al., 2018). Studies of early reading development have reported that relatively few successful exposures to a word are required for the acquisition of word-specific representations in typically developing readers (e.g., Brooks, 1977; Ehri & Saltmarsh, 1995; Reitsma, 1983), implying that word-specific representations form relatively rapidly as children develop reading skills. Furthermore, item-based acquisition acknowledges 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 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 necessarily sufficient to guarantee the formation of a particular word-specific representation (see Nation & Castles, 2017; Share, 2008). Certainly, at the level of the item (e.g., 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 phonology from decoding with the lexical representation of the word (see Share, 2008; Venezky, 1999). Thus, the decodability of a word depends on both the decoding knowledge of the reader and the regularity of the orthographic-to-phonological relationships of the word. 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), might allow a child to determine 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). 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 being influenced by a combination of the reader'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 the overall number of word exposures the child experiences (see Cunningham & Stanovich, 1998; Nation & Castles, 2017; Reitsma, 1983).

Thus, orthographic learning is heavily dependent on a child's ability to go from a decoded form of a word to the stored phonological representation. The ease by which a child can 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 rather awkward term "Set for Variability" is rooted historically in writings by Gibson (1965) and later Venezky (1999), who advocated for the use of phonics instruction for developing readers in English with the important caveat that in order for children to be successful with this approach, they would need a "set for variability" in English. As Venezky put it, "if what is first produced does not sound like something already known from listening, a child has to change one or more of the sound associations (most probably a vowel) and try again" (p. 232). Elbro et al. (2012) offered that this skill serves as a bridge between decoding and lexical pronunciations and may be an important second step in the decoding process.

SfV has been operationalized into an oral language task (see Tunmer & Chapman, 1998) requiring an individual to disambiguate the mismatch between the decoded form of an irregular word and its actual lexical pronunciation. For example, in the task the word *wasp* is pronounced to rhyme with *clasp* (i.e. /*wæsp*/), and the individual must recognize the actual pronunciation of the word to be /wɑsp/. SfV has been reported to explain significant and unique variance in developmental word and nonword reading skill above and beyond that associated with phonemic awareness skill, vocabulary, and letter-sound knowledge (Kearns et al., 2016; Steacy, Compton, et al., 2019; Tunmer & Chapman, 2012). In addition, Tunmer and Chapman (2012) reported that the influence of vocabulary on future decoding and word recognition is mediated by SfV skill. Further, the relationship between word reading development and SfV has been shown to exist across both regular (Kearns et al., 2016) and irregular words (Steacy, Wade-Woolley, et al., 2019) and deep (i.e. Danish) and shallow (i.e. Dutch) orthographies (Elbro et al., 2012). SfV has also been shown to be both an item-specific predictor and general predictor of word reading variance (Steacy, Wade-Woolley et al., 2019), suggesting that the ability to complete the SfV mispronunciation task is important at both the level of the word and the individual. At the level of the individual, Steacy, Wade-Woolley, et al. (2019) reported correlations between SfV and measures of vocabulary, phonemic awareness, and timed word reading of .40, .47, and .77, respectively. In the Steacy et al. study, the relationship between SfV and word reading was strong enough that once SfV was entered into a prediction model, phonemic awareness and vocabulary were no longer significant predictors of item-level word reading.

Given the strength of SfV to predict word reading as an oral language task without print, a need arises for a greater understanding of what contributes to this task. Currently little is known about the word- and child-level demands associated with correctly performing the SfV task. From the perspective of the child, 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). Certainly, the task of disambiguating the mispronunciation to identify the lexical form of the word places primary demands on the phonological system in what can be thought of as a "clean up" 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 contributes to decoding by "cleaning up" noisy or incomplete phonological representations. This allows the hidden units to provide the phonological system with somewhat incomplete input, akin to some type of decoding ambiguity in a child that captures the regularities in the mapping from orthography to phonology but can be easily overridden by the clean-up process in the case of exception words. 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). Similarly, here we surmise that the SfV tasks taps into a similar phonologically-based "clean-up" process in children that allows them to disambiguate the mismatch between the decoded form of a word and its actual lexical pronunciation. It is also quite feasible that a child's SfV performance is affected by characteristics of the items as well. For instance, variation in word features such as the number of phonological competitors to the target word could affect the ease or difficulty of SfV item performance. Thus, the role of phonology, both at the child- and word-level, would appear crucial for completing the task. However, there may be other sources that influence performance on the SfV task in developing readers.

On the child side, Steacy, Compton, et al. (2019) speculated 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. 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. In our case, the process works in the opposite direction such that hearing a regularized form of an irregular word generates the correct orthographic form of the word that in turn aids in the identification of the target word in the SfV task. Both of these hypotheses suggest that there may be orthographic involvement during the SfV task, associated with decoding skill and reading experience, which links task performance to word reading and decoding skill above and beyond the level of straight phonological and semantic processing. In this study, we specifically examine the role of orthographic involvement by examining the relative importance of child decoding (tapping a child's knowledge of phonology to orthography connections) and word reading (a proxy of a child's experience reading words) skills in predicting SfV performance in models that also include phonological measures.

On the word side, very little consideration has been given to word characteristics that might affect SfV performance. Steacy, Compton, et al. (2019) reported that when children were asked to read the SfV items, an expert rating of relative transparency of each word (i.e. spelling to pronunciation transparency rating) was a significant predictor of performance above and beyond measures of word length, frequency, and concreteness, suggesting that it is more difficult for developing readers to go from the decoded form to the stored phonology as spelling to pronunciation transparency decreases (see Waters et al., 1984 for similar findings of the impact of spelling-sound correspondences at the child level). Word-level characteristics have also been reported to predict word difficulty on other reading related tasks such as phonemic awareness (Hogan et al., 2011) and rapid automatized naming (Compton, 2003). In this study, we are interested in whether a measure tapping word-level phonology to orthography consistency of the mispronounced stimulus (i.e. P→O surprisal from the mispronunciation to the target's orthographic form) adds variance above and beyond other word-level characteristics that are more dependent on item phonology (e.g., number of phonological competitors). The assumption here is that it is easier for children to involve orthography during the task on items where the target's orthographic form is more strongly related to the mispronounced spoken stimulus.

The present study is the first to examine word features and child characteristics that contribute to individual differences in item-level performance on the SfV task in a large sample of developing readers. Of particular interest in the current study is whether only predictors that involve phonology explain the variance in SfV or whether predictors that involve the connection between phonology and orthography explain additional variance. Since the mispronunciation task contains no visual written components (orthography) in its administration, it is possible that the task is purely phonological (and semantic) in nature.[1] However, given the strong correlation with word reading, it is possible that SfV task performance is also affected by a child's knowledge of phonology to orthography associations and that orthographic facilitation is more likely in items that have greater consistency between phonology and orthography. Furthermore, we were interested in the role of word reading skill in explaining additional variance and the extent to which it moderates the effects of child-level decoding skill and word-level P→O surprisal on SfV performance. Specifically, we examine whether orthographic facilitation during the SfV task is more likely in children with better word reading and decoding skills through a hypothesized mechanism in which knowledge of connections between phonology and orthography facilitates SfV performance.

In order to investigate these questions, a series of models were constructed to determine the influence of child characteristics and word features on item-level SfV performance by first adding child- and word-level predictors representing general control variables

---

[1] In this study we control for child vocabulary skill as a general predictor and expressly examine the role of phonology only versus combined phonology and connections between phonology and orthography as predictors of SfV variance associated with child and word.

(i.e. child rapid naming skill, vocabulary knowledge, and attention rating; word frequency, length, concreteness, and order of administration), next adding variables that tap phonology (i.e., number of phonological competitors for the word and child phonemic awareness skill), followed by adding variables that tap phonology-to-orthography connections (i.e., word P→O surprisal and child word attack skill), and finally adding child word reading skill to explore its role as a moderator between child- and word-level predictors. We included appropriate interactions at various points in the model building process to help better understand the relationships between child- and word-level features as predictors. We expect orthographic involvement in SfV performance, hypothesizing that both child- and word-level variables tapping phonological to orthographic connections will uniquely predict SfV performance after accounting for associated tasks tapping only child- and word-level phonological skill, but only in children with sufficient decoding skills to activate phonological to orthographic connections. In addition, we predict that word identification skill, considered a child-level predictor of experience reading words, will moderate the relationship between child- and word-level features related to phonology to orthography connections and SfV performance.

Method

Participants

Participants consisted of 489 children in grades 2–5 from 5 local public schools in the southeast and 4 private schools serving children with learning differences located along the east coast of the United States. This sampling plan oversampled for children struggling to learn to read. However, standard scores on norm-referenced measures of phonemic awareness, rapid letter naming, vocabulary, word identification, and word attack (8.67 (*SD* = 2.91), 9.08 (*SD* = 2.21), 10.08 (*SD* = 3.14), 100.91 (*SD* = 13.42), 100.48 (*SD* = 10.34), respectively) indicate a sample that performed within the normal range on national norms. Sample demographic data broken down by school type is provided in Table 1.

Table 1

Demographic Breakdown by School Type

Demographic	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 word attack SS	98.0 (9.7)	101.8 (10.4)
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. ELL = English language learner; SS = standard score; LD = learning disability; WJ = Woodcock-Johnson; WASI = Wechsler Abbreviated Scale of Intelligence; CTOPP = Comprehensive test of Phonological Processing. Standard deviations are in the parentheses.

Procedure

Children were assessed one on one in schools by trained research assistants who underwent extensive training and were required to achieve 80% procedural fidelity before testing participants. Testing sessions were audio recorded for scoring and fidelity purposes. All tests were double scored and double entered by a fellow research assistant, with any discrepancies being resolved by the project coordinator. Average fidelity of test administration (based on a random selection of 20% of the taped assessment sessions) was completed, with procedural fidelity and interrater agreement exceeding 98% across all tests.

Child Measures

Set for Variability (SfV)

SfV was measured using a mispronunciation correction task in which participants heard a recording of a word pronounced incorrectly and were asked to respond verbally with the correct pronunciation. Participants were told that they were going to play a word game with Alex and that they needed to try to figure out what word Alex is trying to say. They were given two practice items ("/mŏther/" and "/brēkfăst/") with corrective feedback before they began the task, followed by an additional 75 test items based on the original Tunmer and Chapman (1998) measure. These mispronunciations consisted of a regularized pronunciation (i.e. decoded form) of each irregular word, which was derived by sounding out a word as it could be pronounced when applying decoding rules using the highest frequency grapheme-phoneme correspondences and most frequent syllable patterns. In some cases, the mispronunciations resulted in phoneme substitutions (e.g., the target word *breakfast* was mispronounced as */brik fəst/*) and in others it resulted in phoneme insertions (e.g., the target word *lamb* was mispronounced as */læmb/*such that the “b” was articulated rather than silent). Performance ranged from 1 to 58. Ordinal alpha for this sample was .95.

Phonological Awareness

Phonological Awareness was measured by the Elision task from the Comprehensive test of Phonological Processing 2 (CTOPP−2; Wagner et al., 2013). Students were asked to delete phonological units from words. In this sample, performance ranged from 0 to 34. The authors report test–retest reliability of .93 (Wagner et al., 2013).

Word Identification

Word identification was measured by the Letter–Word Identification subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock et al., 2001), an untimed measure of word reading ability. Performance ranged from 23 to 67 in this sample. The authors report a split-half reliability of .93-.96 for the ages assessed here (McGrew & Woodcock, 2001).

Word Attack

Word Attack was measured by the Word Attack subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock et al., 2001), an untimed measure of nonword reading ability. In this study the word attack task is used as a proxy for children's knowledge of connections between phonology and orthography. Raw scores ranged from 4 to 30 in this sample. The authors report a split-half reliability of .88-.92 for the ages assessed here (McGrew & Woodcock, 2001).

Rapid Letter Naming

Rapid Letter Naming was assessed using the Rapid Letter Naming subtest of the CTOPP−2 (Wagner et al., 2013). Participants named letters in serial order as quickly and accurately as they could. Scores indicated the number of seconds taken to name all letters, ranging from 11 to 59 in this sample. Test–retest reliability is .72 for children of ages 8–17 years according to the test manual (Wagner et al., 2013).

Vocabulary

Vocabulary was assessed using the vocabulary subtest of the WASI (Wechsler, 2011), a measure of expressive vocabulary. The test requires students to identify pictures and define words. Performance ranged from 1 to 39 in this sample. Interrater reliability for elementary age children ranges from .92–.94 (McCrimmon & Smith, 2013).

Inattention

Inattention was measured by the inattention items of the Strengths and Weaknesses of ADHD-Symptoms and Normal Behavior Scale (SWAN; Swanson et al., 2012). The inattention score consists of 9 items filled out by teachers based on their observations of the child. Scores are scaled such that higher scores indicate more inattention behaviors. Scores ranged from 9 to 63 in this sample. Ordinal alpha for these items in this sample was .96.

Word Measures

Note that when we refer to the "stimulus," this is the provided oral mispronunciation on the SfV task, whereas the "target" refers to the correct response.

Length

This is the number of letters in the target word, ranging from 4 to 8 in these items.

Frequency

We used the log transformed HAL frequency of the target word taken from the English Lexicon Project (Balota et al., 2007) to represent frequency. This measure of frequency is based on the Hyperspace Analogue to Language corpus, which consists of approximately 131 million words gathered from 3,000 newsgroups during 1995. Frequency values for items in this study ranged from 6.4 to 12.53.

Order

The order in which stimulus items were presented during administration, one being the first item presented to students and 75 being the last. Each child heard the items in the same order, with the order being set randomly (i.e. items were not ordered by difficulty). While this predictor is not strictly considered a "word-level" feature, it is associated with word-level variance, and therefore we refer to it as a word-level predictor.

Concreteness

Concreteness of the target was represented by ratings taken from the English Lexicon Project provided by Brysbaert et al. (2014). Three words did not have ratings so the rating for the singular version of each word was used (i.e., for *heights, muscles*, and *veins* ratings were taken from *height, muscle*, and *vein*). Items here ranged from 1.55 to 5 on concreteness.

Number of Competitors

We calculated the number of competitors based on phonological Levenshtein distance. Phonological Levenshtein distance is the phonological distance between the stimulus (i.e. the mispronunciation) and target (i.e. the actual pronunciation of the word), which is operationalized as the minimum number of single phoneme edits (i.e. insertions, deletions, or substitutions) required to change the stimulus into the target. Thirty-four items had a phonological distance between stimulus and target of 1; 27 items had a distance of 2; 9 items had a distance of 3; and 5 items had a distance of 4. Number of competitors represents the number of words that can be made by applying edits of the Levenshtein distance between the stimulus and target or fewer edits (i.e. how many words are as close or closer to the stimulus than the target). This measure is log transformed to correct for skewness. The log transformed number of competitors and Levenshtein distance were highly correlated at .90, thus only number of competitors was included as a predictor in the models. Since the number of competitors is determined only by the phonological distances between the stimulus, target, and possible competitors, we treat it as our phonology-related metric in the models below, closely related to the notion of a "phonological neighborhood" (Vitevitch & Luce, 2016). Items here ranged from 1.55 to 5 on concreteness.

Phonological-Orthographic Consistency (Max P→O Surprisal)

This measure represents how likely it would be for the spelling of a word to be derived from the mispronunciation. Specifically, higher values represent a word that contains a phoneme whose associated spelling is less likely for that phoneme. For each phoneme in each stimulus (i.e. the mispronounced spoken form), we calculated the conditional probability of the corresponding grapheme given the (mispronounced) phoneme (i.e. p[grapheme. Probabilities were calculated using the phonological database by Kearns (2020); a corpus of 117,574 English words and their pronunciations. Then, probability estimates were transformed into surprisal values (i.e., $-log(p(i))$); thus, higher values represent more unlikely pairings (see Siegelman et al., 2020 for a discussion of the utility of information-theoretic measures in capturing orthographic-phonological associations). As a result of this procedure, each stimulus had K surprisal values, where K is the word length in phonemes/graphemes. From this series of values, we use as an item-level predictor in the models below the maximal surprisal value.[2] This variable was log-transformed to account for skewness in the data. This measure presumably reflects the extent of difficulty in "translating" a word's mispronounced spoken form to its orthographic form. For example, the item with the lowest surprisal was *lamb*, which was pronounced */læmb/*. Each phoneme in the mispronunciation was represented by a grapheme that is highly likely to orthographically represent that sound. Whereas in "rhythm" the phoneme /r/ was represented by the grapheme "rh", which is an unlikely representation of that phoneme. Values ranged from .09 to 13.84 prior to transformation and −2.41 to 2.63 after.

Analytic Plan

Cross-Classified Generalized Random-Effects Models

SfV item responses were modeled using cross-classified generalized random-effects (CCGRE) models using the *lme4* package in R (Bates et al., 2015). These models allow for the estimation of variability in item responses between children as well as between words. In these models, children are crossed with words, and both children and words are allowed to be random factors. Since all of the responses were dichotomous (correct/incorrect), a binomial distribution with a logit link function was used to predict the probability of a correct item response based on the set of predictors. For all models, variables were z-scored to aid in interpretation such that 0 represents the mean, and a value of 1 represents 1 standard deviation above the mean. In order to estimate the additional variance accounted for by phonological predictors, predictors representing the connection between phonology and orthography, and word reading skill, a series of 5 models were conducted: unconditional model, general predictors, phonological predictors, phonology-orthography connection predictors, and a full prediction model with the inclusion of word identification. Appropriate interactions between child characteristics and word features were included in later models. After controlling for age, the intraclass correlation was near zero for teacher and was .22 for school. To account for the nesting structure in which children were nested within schools, a random effect for school was included in all models.

Model Building and Interpretations

The unconditional model (Model 0) provides the total amount of variance in SfV performance at both the child and word level. This is used to determine the amount of variance explained by predictors in the subsequent models. The first predictors in Model 1 are termed general predictors. These predictors are used to account for variance in SfV performance that can be explained by other associated measures not specific to phonology or phonology-orthography connections. Then, predictors representing phonology are added to the model in Model 2. Phonological awareness and number of phonological competitors both involve phonology without involving orthography (i.e. print/spellings are not involved). Additional variance accounted for by this model that was not explained by the general predictor model represents the unique contribution of phonology in SfV. The next model (Model 3) adds predictors that represent the connection between phonology and orthography. At the word level, max P→O surprisal represents how well the mispronunciation matches the spelling of the word. At the child level, word attack (WA) is used as a proxy for knowledge of the connection between phonology and orthography. Performance on the WA task is heavily contingent on a child's ability to go from orthography to phonology using knowledge of decoding rules without the influence of semantics given that items are nonwords. After controlling for phonological knowledge in the previous model, WA should provide an indication of the extent to which knowledge of the connection between orthography and phonology predict SfV performance. While WA is not a pure indicator of knowledge of the connection between orthography and phonology, with additional variance likely unrelated to knowledge of the connection between orthography and phonology, WA after controlling for phonological awareness is used as a proxy for this connection knowledge at the level of the child. The additional variance explained in Model 3 represents the contribution of the knowledge of the connection between phonology and orthography after accounting for the influence of phonology alone. Lastly, word identification is added to the model in Model 4. We conceptualize word identification performance as a proxy for experience reading words, which allows children to build up word-specific knowledge along with general relations between orthography and phonology (see Perfetti, 1992), and thus it is added to the model last. While the additional variance directly attributable to word identification is not uninteresting, we are more interested in whether word reading skill acts as a moderator of the effects of child and word measures tapping connections between orthography and phonology and SfV performance. This allows us to test whether having greater reading skill increases the probability that children will use orthographic information to perform the SfV task and whether that probability increases as the consistency of SfV items increases. Specifically, we include three-way interactions with word identification between word attack and word-level features (number of phonological competitors and max P→O surprisal) observed in Model 3 in order to determine whether these relations are due solely to decoding skill or whether these relations also depend on a child's level of word reading skill. Furthermore, when word identification and word attack are included in an interaction together it, helps to disentangle whether SfV is impacted by decoding skill, process(es) involved in word reading unshared with decoding skills, or both. This provides a greater understanding of the underlying skills that contribute to successful completion of the SfV task.

Results

Descriptive statistics and correlations for the child-level predictors are depicted in Table 2. Because of the relatively wide age range of the sample we provide zero-order correlations below the diagonal and age-corrected correlations above the diagonal. All child-level predictors were significantly correlated with SfV, with large associations with word identification (.79), word attack (.76), phonological awareness (.66), and vocabulary (.50), and moderate associations with rapid letter naming (−.35) and inattention (.33). In general, correlations between age and the other child-level predictors were low to moderate, and controlling for age had little effect on the correlational matrix between child-level predictors and

---

[2] We opted to use maximum rather than mean surprisal across phoneme-grapheme pairings because in most stimuli there were only one or two mispronounced phonemes (typically vowels), and hence surprisal in most pairings was zero (e.g. correctly read consonants with no alternative spellings).

Table 2

Means, Standard Deviations, and Zero-Order and Age-Corrected Correlations for Child Level Predictors

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. Age	9.17	1.13								
2. RLN	20.63	5.65	−.22**		−.16**	−.19**	−.22**	−.28**	−.41**	−.31**
3. Vocabulary	23.86	6.33	.32**	−.22**		.32**	.46**	.36**	.40**	.46**
4. Inattention	37.57	11.45	−.13**	−.15**	.26**		.34**	.40**	.45**	.38**
5. PA	22.17	6.72	.14**	−.24**	.48**	.31**		.69**	.60**	.65**
6. Word attack	17.94	6.11	.17**	−.28**	.39**	.37**	.70**		.78**	.75**
7. WID	45.97	8.20	.30**	−.43**	.46**	.38**	.61**	.78**		.78**
8. SfV	26.74	11.17	.26**	−.35**	.50**	.33**	.66**	.76**	.79**	

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. RLN = rapid letter naming; PA = phonological awareness; WID = word identification; SfV = set for variability. Zero-order correlations are below the diagonal; age-corrected correlations are above the diagonal.
** $p < .01$.

SfV. The correlation between SfV and word identification and word attack was consistent with (Steacy, Compton, et al., 2019; Steacy, Wade-Woolley, et al., 2019) but somewhat higher than those reported by Elbro et al. (2012) and Kearns et al. (2016). Table 3 shows descriptive statistics and correlations for the item-level predictors. Significant associations were observed between SfV and length (.40), order (−.34), and number of competitors (−.43), whereas correlations with frequency, concreteness, and max P→O surprisal were not significant.

Unconditional Model

First, an unconditional model was fit to estimate the variance in item responses attributable to children, variance attributable to words, variance attributable to school, and a grand mean (intercept) that yields the probability of getting an item correct (in logits). Results revealed both variability due to children (*SD* = .87) and variability due to words (*SD* = 1.61), indicating enough variability to attempt to predict this variability with our predictors. Although no school-level predictors are investigated here, the random effect for school (*SD* = .41) is included to account for the nesting structure present in the data. Results showed that on average across children and words, the probability of a correct response was .27 [intercept = −.98, z = −4.16, $p < .001$]. Although this shows fairly low performance overall, there was significant spread in the distribution with no floor effects present in the data. The total correct on the SfV task was normally distributed (Skewness = .389 (.108), Kurtosis = −.326 (.215)) with no child getting none correct.

General Predictor Model

Next, general predictors were added to the model to examine the influence of word features (i.e. number of letters, frequency, order in administration, and concreteness) and child characteristics (i.e. age, rapid letter naming, vocabulary, and inattention) that neither represent phonology nor the connection between phonology and orthography.[3] Results are reported in Table 4 (Model 1). All predictors contributed significantly to the prediction except for the order in which items were presented in task administration. These general predictors accounted for 36.61% of the child level variance, 30.43% of the word level variance, and 50.18% of the school level variance in SfV performance.

Phonological Predictor Model

Next, both word (i.e. number of phonological competitors) and child phonological predictors (i.e. phonological awareness) were added to the model, alongside the predictors from the previous model to control for the variance these predictors already explain. An interaction term was included between phonological awareness and number of competitors to investigate whether the influence of the number of phonological competitors depends on the child's level of phonological awareness. Results are presented in Table 4 (Model 2). Age, length, frequency, concreteness, rapid letter naming, vocabulary, inattention, number of competitors, phonological awareness, and the interaction between phonological awareness and number of competitors all contributed significantly, whereas item order did not. The interaction between phonological awareness and number of competitors indicates that the impact of number of competitors is greater for those with lower phonological awareness skill. This interaction fails to reach statistical significance in subsequent models with additional predictors. This model explained 55.83% of the child-level variance, 47.79% of the word-level variance, and 76.66% of the school-level variance, an extra 19.22%, 17.36%, and 26.48%, respectively, above the general predictors.

Phonology-Orthography Connection Predictor Model

Next, word and child predictors representing the connection between phonology and orthography (i.e. max phonology to orthography P→O surprisal rating and word attack, respectively) were added to the model. Interaction terms were added between word attack and number of competitors and max P→O surprisal. Results are reported in Table 4 (Model 3). Significant contributions were observed by age, length, frequency, concreteness, rapid letter naming, vocabulary, number of competitors, phonological awareness, word attack, the interaction between number of competitors and word attack, and the interaction between word attack and max P→O surprisal. In contrast, order, inattention, max P→O surprisal and the interaction between number of competitors and

---

[3] We included rapid letter naming as a control variable because it is unclear whether it should be considered a pure measure of phonology or a measure tapping phonology-orthography connections. This decision does not affect parameter estimates since all models rely on simultaneous inclusion of predictors, with the only effect being the relative R-square estimates for models 1, 2, & 3.

Table 3
Means, Standard Deviations, and Correlations for Word Level Predictors

Variable	*M*	*SD*	1	2	3	4	5	6
1. Length	5.47	1.20						
2. Frequency	8.95	1.43	−.40**					
3. Order	38.00	21.79	−.31**	.02				
4. Concreteness	4.11	0.97	.29*	−.26*	.04			
5. Number of competitors	4.36	2.28	−.10	.09	.14	.12		
6. Max P→O surprisal	0.88	1.03	.23*	−.10	−.20	−.18	.25*	
7. Set for variability	0.36	0.24	.40**	.10	−.34**	.20	−.43**	−.07

Note. M and *SD* are used to represent mean and standard deviation, respectively.
* $p < .05$. ** $p < .01$.

phonological awareness did not contribute significantly to the prediction. The interaction between word attack and max P→O surprisal is depicted in Figure 1 and shows that those with better decoding skills are impacted by greater max P→O surprisal (note that the analyses are conducted in logits, and these graphs are converted to probability. This can create a nonlinear graph due to the approaching of 0). The interaction between word attack and number of competitors indicates that the probability of a correct response is lower when the number of competitors is large, especially for those with low word attack skills. This model explained 69.44% of the child-level variance, 48.69% of the word-level variance, and 92.18% of the school-level variance, an extra 13.61%, .90%, and 15.52%, respectively, above the phonological predictor model.

Full Prediction Model

Lastly, word identification was added to the model to determine whether additional variance is explained by word reading skill beyond decoding skill and phonological awareness. A number of interaction terms were added that included: two-way interactions between word identification and word attack; word identification and number of competitors; word identification and max P→O surprisal; three-way interactions between word identification,

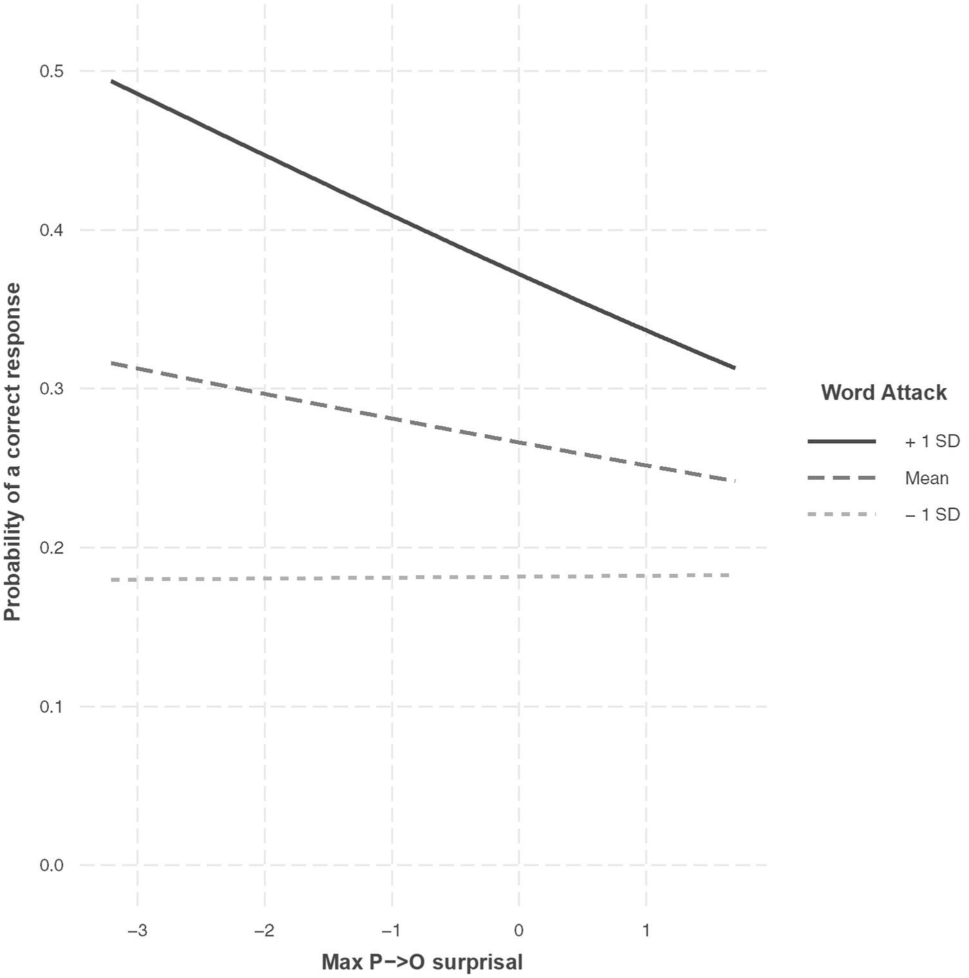
Table 4
Results of the Predictor Models (1–4)

	Model 1			Model 2			Model 3			Model 4		
Parameter	Logit	z-value	*OR*	Logit	z-value	*OR*	Logit	z-value	*OR*	Logit	z-value	*OR*
Intercept	−1.04	−5.49**	0.35	−1.02	−6.61**	0.36	−1.01	−7.10**	−1.01	−1.03	−7.52**	0.36
General predictors												
Age[b]	0.20	4.57**	1.22	0.18	4.81**	1.20	0.11	3.33**	1.65	0.02	0.85	1.02
Length[a]	0.57	3.10**	1.77	0.53	3.38**	1.70	0.55	3.34**	1.73	0.55	3.34**	1.73
Frequency[a]	0.53	3.05**	1.70	0.60	3.97**	1.82	0.59	3.93**	1.80	0.59	3.94**	1.80
Order[a]	−0.32	−1.95	0.73	−0.25	−1.74	0.78	−0.27	−1.82	0.76	−0.26	−1.82	0.77
Concreteness[a]	0.43	2.56*	1.54	0.54	3.64**	1.72	0.51	3.30**	1.67	0.51	3.31**	1.67
RLN[b]	−0.19	−5.11**	0.83	−0.14	−4.16**	0.87	−0.10	−3.56**	0.91	−0.04	−1.36	0.96
Vocabulary[b]	0.30	6.84**	1.35	0.17	4.33**	1.19	0.17	4.96**	1.16	0.13	3.20**	1.14
Inattention[b]	0.21	5.52**	1.23	0.11	3.33**	1.12	0.04	1.20	1.04	0.00	0.08	1.00
Phonological predictors												
NC[a]				−0.69	−4.93**	0.50	−0.67	−4.53**	0.51	−0.69	−4.69**	0.50
PA[b]				0.47	13.14**	1.60	0.19	4.86**	1.21	0.17	4.40**	1.19
NC x PA[ab]				0.06	3.83**	1.06	0.00	−0.11	1.00	−0.01	−0.54	0.99
P→O predictors												
WA[b]							0.49	12.62**	1.70	0.26	5.74**	1.30
P→O S[a]							−0.08	−0.49	0.92	−0.05	−0.31	0.95
NC x WA[ab]							0.11	5.49**	1.12	0.04	1.72	1.04
WA x P→O S[ab]							−0.08	−5.05**	0.92	−0.05	−1.98*	0.95
Additional predictors												
WID[b]										0.41	9.23**	1.51
WID x WA[bb]										0.06	2.79**	1.06
WID x NC[ab]										0.09	3.72**	1.09
WID x P→O S[ab]										−0.04	−1.51	0.96
WID x WA x NC[ab]										0.03	1.86	1.03
WID x WA x P→O S[ab]										−0.03	−2.32*	0.97

[a] Word-level predictor. [b] Child-level predictor. [ab] Word- by child-level interaction. *OR* = odds ratio; RLN = rapid letter naming; NC = number of competitors; PA = phonological awareness; WA = word attack; P→O S = Max P→O surprisal; WID = word identification.
* $p < .05$. ** $p < .01$.

Figure 1

Interaction Between Max P→O Surprisal and Word Attack From the Phonology-Orthography Connection Predictor Model

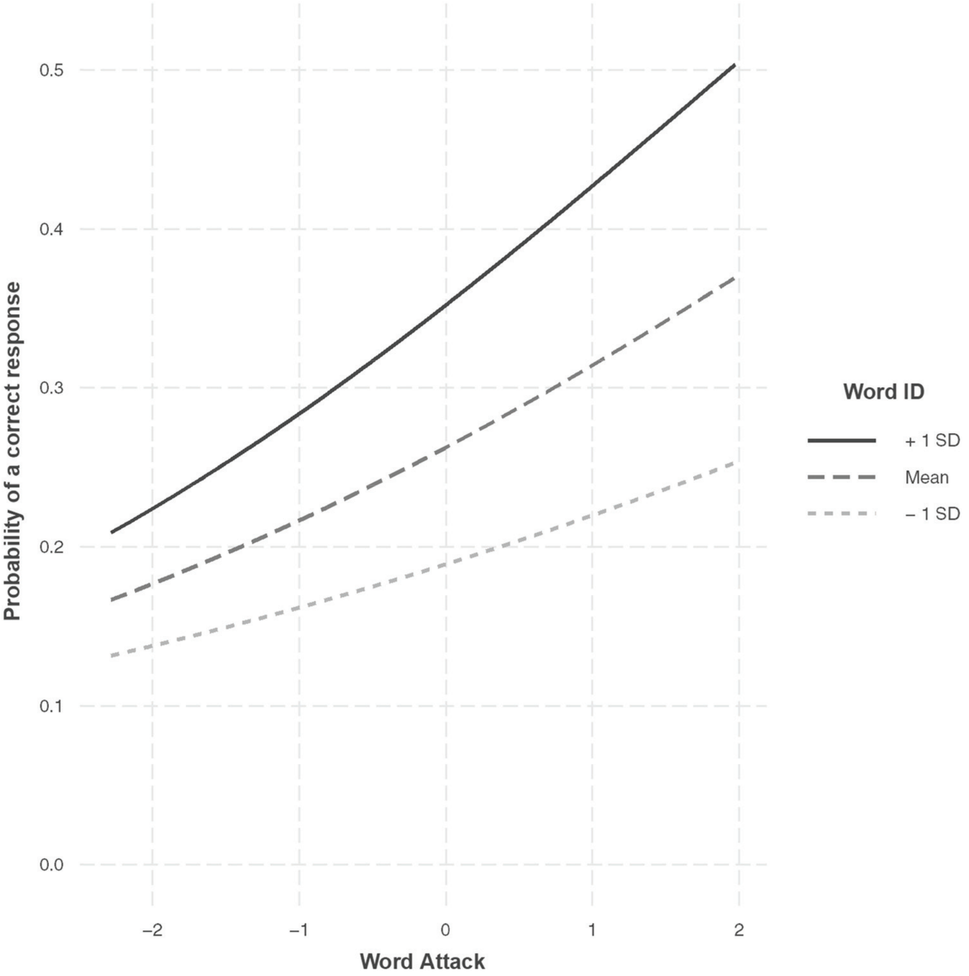


Note. Max P→O surprisal is represented in *z*-score units.

word attack, and number of competitors; and finally word identification, word attack, and max P→O surprisal to determine whether these relations change depending on the level of word reading skill. Results are presented in Table 4 (Model 4). Length, frequency, concreteness, vocabulary, number of competitors, phonological awareness, word attack, word identification, the interaction between word attack and word identification, the interaction between number of competitors and word identification, the interaction between word attack and max P→O surprisal, and the three-way interaction between word attack, word identification, and max P→O surprisal all contributed significantly to the prediction. In contrast, age, order, rapid letter naming, inattention, max P→O surprisal, the interaction between number of competitors and phonological awareness, the interaction between number of competitors and word attack, the interaction between word identification and max P→O surprisal, and the three-way interaction between word attack, word identification, and number of competitors were not significant unique contributions. The main effect of word identification suggests that at the mean of all other predictors, the probability of a correct response is .35 at 1*SD* above the mean on WID compared to a probability of .19 at 1*SD* below the mean on WID. The main effect for word attack suggests that at the mean of all other predictors, the probability of a correct response is .32 at 1*SD* above the mean and .22 at 1*SD* below the mean. For vocabulary, the probability at 1*SD* above the mean was .29 compared to .24 at 1*SD* below the mean and for phonological awareness, the probability was .30 at 1*SD* above the mean compared to .23 at 1*SD* below the mean. Likewise, for word-level predictors, the probability of correctly identifying an SfV item at the mean of all other predictors was .38 at 1*SD* above the mean length and .17 at 1*SD* below the mean on length, meaning that the odds of a correct response are about 3 times higher for longer words (1*SD* above mean) compared to shorter words (1*SD* below the mean). For frequency, the probability was .39 at 1*SD* above the mean and .17 at 1*SD* below the mean. For concreteness, the probability of a correct response was .37 at 1*SD* above the mean on

Figure 2

Interaction Between Word Attack and Word Identification From the Full Prediction Model



Note. Word attack is represented in *z*-score units.

concreteness and .18 at 1*SD* below the mean. Finally, for number of competitors the probability of a correct response was .15 when the number of competitors was 1*SD* above the mean and .42 when the number of competitors was 1*SD* below the mean, meaning that the odds of a correct response were over 4 times higher when words were 1*SD* above the mean on number of competitors compared to words that were 1*SD* below the mean.

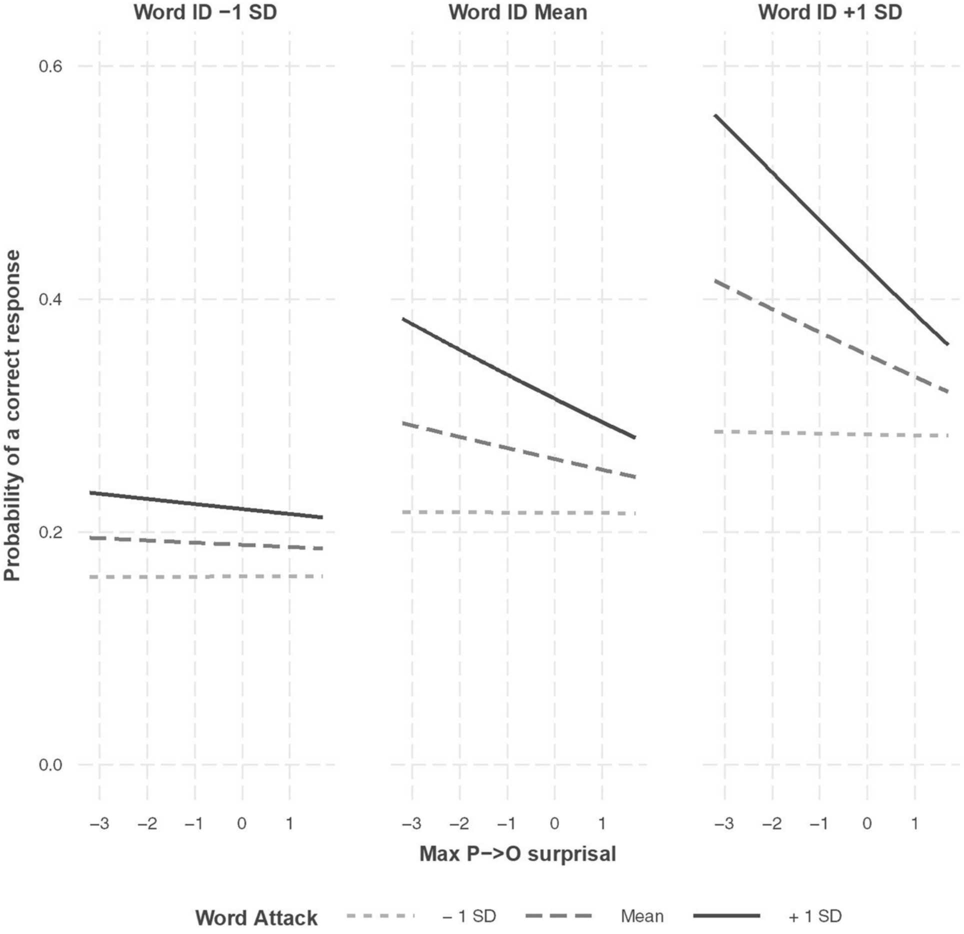
The interaction between word identification and number of competitors implies that number of competitors had a greater impact on those with poor word identification skills. The interaction between word attack and word identification is depicted in Figure 2 and shows that children with high word attack skills perform better on SfV, but even better when children are also high performers on word identification. Furthermore, the three-way interaction between word attack, word identification, and max P→O surprisal (see Figure 3) shows that children high on both word identification and word attack show a greater impact of consistency between phonology and orthography. Interestingly, this interaction also shows that those with low word attack skills are essentially unaffected by this consistency regardless of word identification skills. This suggests that knowledge of the connection between orthography and phonology is required to observe the orthographic facilitation when phonology matches closely with orthography regardless of word reading ability. This model explained 76.01% of the child-level variance, 48.82% of the word-level variance, and 99.17% of the school-level variance, an extra 6.57%, .13%, and 6.99%, respectively, above the phonology-orthography connection predictor model.

Discussion

Expanding our understanding of the underlying mechanisms associated with orthographic learning is key to both describing and influencing early reading development in children (Castles et al., 2018, Rayner et al., 2001). SfV has emerged as a powerful and unique predictor of word reading skill in developing readers, however little is known about the demands associated with correctly performing the task. The purpose of this study was to explore the word features and child characteristics that impact SfV by

Figure 3

Three Way Interaction Between Max P→O Surprisal, Word Attack, and Word Identification From the Full Prediction Model



Note. Max P→O surprisal is represented in *z*-score units.

examining whether orthographic knowledge facilitates performance above and beyond phonological skill, and if so, for whom and for which word types. This was accomplished by building a set of hierarchical models that allowed us to examine, at the level of both child and word, the unique contributions of predictors representing phonology only (i.e. phonemic awareness and number of phonological competitors) and the connection between phonology and orthography (i.e. word attack and P→O surprisal). In addition, we built a final model to test word identification skill as a moderator of the effects of word and child predictors on SfV performance.

Overall, we were able to predict a substantial amount of the item-level variance associated with word features (48.82%) and child characteristics (76.01%). Unique word features that predicted SfV variance included word length, frequency, concreteness, and number of phonological competitors (negative coefficient), suggesting that longer, more frequent, concrete words, that are somewhat unique phonologically are easier for children to disambiguate. On the child side, unique predictors included vocabulary skill, phonological awareness, word attack, and word identification, further supporting the strength of the relationship between SfV and reading development, decoding, vocabulary, and phonological skills. We structure our discussion by presenting the results of child characteristic predictors, then word feature predictors, and finally the interactions between child characteristics and word features as influencers of SfV item performance. We chose to emphasize results in this order because while the results of child- and word-level predictors are interesting on their own, it is really the interactions between word- and child-level predictors that allow us to evaluate the hypothesis that knowledge of connections between phonology and orthography facilitates SfV performance.

Effects of Child Characteristics

Results indicate that the addition of child predictors representing the connection between phonology and orthography accounted for additional variance at the child level above and beyond what was explained by phonological and general predictors. The significant main effect of both word attack and word identification when entered simultaneously in the model indicates unique variance explained by each skill. Consistent with the predictions made by Steacy, Compton, et al. (2019), performance on the mispronunciation task relies on shared processes with word reading above those also shared with decoding, phonological awareness, and vocabulary skills. Results suggest that children with better word identification skill show an additional advantage on SfV even after controlling for their word attack skills and general control variables, suggesting that there are skills uniquely shared between word identification and SfV not shared with word attack. Given that reading real words requires the extra step of matching the decoded form of the word to the stored pronunciation (a step that is not present in reading of nonwords), it is likely that what is being represented by this unique variance explained by word identification above word attack is attributable to one's ability to be flexible with this connection. However, an alternate explanation, based on a more unitary conceptualization of word reading development, posits that better word readers have a stronger knowledge of both specific word spellings and the sublexical connections between orthography and phonology that may be accessible to aid in mispronunciation correction. Further, a unitary model argues that during reading development more skilled word readers are better able to build knowledge of the probabilistic constraints between orthography and phonology that are likely necessary to support orthography facilitation of irregular words that make up the SfV task (see Compton et al., 2014; Ehri, 2014; Perfetti, 1992).

The significant interaction between word attack and word identification further clarifies the relationship between word attack, word identification, and SfV performance. The impact of word identification skill was greater when word attack skill was high. This suggests that those with better decoding skills are better able to use this unique contribution of word identification skills (whether that be flexibility or strength of sublexical representations), whereas those with poor decoding skills are less impacted by their word reading skill. This supports the primacy of decoding skill in allowing orthographic facilitation and also the possibility of unique processes shared between SfV and word reading unshared with decoding skill.

Effects of Word Features

On the word side, the addition of predictors representing the connection between phonology and orthography did not account for substantial additional variance at the word level above and beyond what was explained by phonological and general predictors. The number of phonological competitors was a unique predictor with SfV performance decreasing as the number of phonological competitors between stimulus and target increased. Presumably the activation of multiple phonological competitors makes it more difficult to isolate the correct word. This was evident on target words such as *sugar*, pronounced as /*su-gɑr*/on the SfV task, in which many children incorrectly identified it as the target word *cigar*. P→O surprisal, on the other hand, was not a significant predictor of SfV item performance across participants when in the presence of other predictors. Presumably, for P→O surprisal to be a significant predictor of SfV performance, suggestive of orthographic facilitation, developing readers would not only need to have knowledge of phonological to orthographic connections but also be able to activate this knowledge during the SfV task. As we will argue in the next section, being able to activate and use phonological to orthographic connections during the task likely requires a certain level of decoding and word reading skill.

Interactions Between Child Characteristics and Word Features

The interactions between child characteristics and word features allow for important insights regarding whether orthographic knowledge facilitates SfV performance, and if so, for whom and for which words. The interpretation of the main effect of word attack as supporting the hypothesis that knowledge of the connection between orthography and phonology is important for successful completion of the SfV task has certain shortcomings, least of which is that word attack is not a pure indicator of the connection between orthography and phonology. However, the interaction between word attack and max P→O surprisal provides additional support for the idea that it may in fact be knowledge of connections between phonology and orthography that contributes to successful completion of the SfV task. The max P→O surprisal measure is based on the probability of each grapheme in the target given the presented phoneme in a stimulus, with words higher in P→O surprisal containing less consistent relationships between phonology and orthography. Sensitivity to sublexical relationships between graphemes and phonemes on the child's part would appear essential for children to be sensitive to variations in phonological to orthographic consistency across SfV items. Indeed, we observed that the association between word attack and SfV accuracy was most pronounced for low P→O surprisal (i.e. high phoneme-grapheme consistency) words, suggesting that those children with better decoding skills may be using their knowledge of the varied connections between phonology and orthography to aid in pronunciation correction whereas children with poor decoding skills may be relying on phonology only. This is consistent with the idea 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.

Results of the three-way interaction between word attack, word identification, and max P→O surprisal further support the idea that better decoders are more likely to use orthographic information to complete SfV items. Results of the interaction showed that regardless of word reading skill, those low in decoding skills were unable to use consistency between phonology and orthography to aid in mispronunciation correction whereas those with both high decoding and word reading skill showed the greatest impact of this orthographic facilitation. The fact that word reading skill moderated the relationship between word attack and P→O surprisal, but only at higher levels of word attack skill, seems to support the primacy of decoding skill in supporting orthographic facilitation during SfV task performance.

We interpret the results as generally consistent with the "mapping hypothesis" explanation of poor decoding skill development (Harm et al., 2003). The mapping hypothesis states that children with poor phonological representations tend to rely on more holistic mappings from orthography to phonology that result in poor generalization across words at the level of sublexical grapheme-phoneme connection (for details see Steacy et al., 2021). As our results suggest, poor word reading and, in particular, poor decoding skill seem to limit children's ability to activate orthography to improve SfV performance. Thus, despite the fact that the SfV task is an oral language task in which students are presented with no print, performance on the task increases when the child is able to use knowledge of the connection between phonology and orthography and when words have greater consistency between phonology and orthography. We therefore argue that SfV taps more than simply knowledge of phonology, which explains why SfV adds unique variance to the prediction of word reading above phonological awareness and vocabulary skill. Additionally, the unique contribution of word identification above and beyond word attack suggests that there are unique processes shared between word identification and SfV not shared with word attack. This suggests that even when a student possesses adequate decoding skills to arrive at the decoded form of the word, the ability to make the jump from this decoded form to the actual pronunciation is important for word reading.

Overall, this study was designed to help better understand what child characteristics and word features contribute to SfV performance. Clearly both types of features are important, and in particular the interactions between child- and word-level predictors shed important light on the relationship between reading skill and SfV performance. Results suggest that knowledge of the connection between phonology and orthography is important in mispronunciation correction, but that decoding ability contributes uniquely as well and that students who possess both decoding and word reading skills perform better on SfV than students possessing only one or the other skill. Understanding what contributes to SfV performance helps to provide insight into why it is such a good predictor of word reading. Furthermore, understanding the word features that contribute to item difficulty is important for selecting items for SfV assessment as well as providing insight into how students are impacted by these features differently based on their word reading and decoding skills. This study also indicates that those with knowledge of the connection between phonology and orthography are able to use this knowledge to benefit from consistency between phonology and orthography, with this orthographic facilitation only occurring when students possess adequate decoding skills.

Clearly, more work is needed to further understand the child and word features that contribute to the task. For instance, the role of child semantic knowledge was not explored in the current study except to control for child (vocabulary) and word (concreteness) variance as control variables. As far as words are concerned, little is known about how actual mispronunciations contribute to SfV item difficulty. For instance, does item difficulty decrease if a decoded form is used versus some other mispronunciation (e.g., a mispronunciation that is not consistent with the word's orthography)? We also wonder how spelling might play into individual differences in SfV performance. Certainly, this study serves as a strong start to exploring important child characteristics and word features that predict SfV performance, but many important questions still require study. Finally, there is emerging evidence that lexical flexibility, based on the SfV task, can be trained in children (Dyson et al., 2017; Savage et al., 2018; Zipke, 2016). Training protocols have emphasized flexibility in applying different pronunciations for letters or letter combinations (Zipke, 2016), checking for matches and making approximations to known words (Dyson et al., 2017; Savage et al., 2018), and a two-step instructional model where direct instruction in simple decoding was the first step and set for variability flexibility training followed as a second step (Savage et al., 2018). A better understanding of the SfV task has the potential to better inform these training efforts.

References

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459. https://doi.org/10.3758/BF03193014

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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). Erlbaum.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 151–179). Psychology Press.

Compton, D. L. (2003). The influence of item composition on RAN letter performance in first-grade children. *The Journal of Special Education*, *37*(2), 81–94. https://doi.org/10.1177/00224669030370020301

Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of "quick fix. *Interventions for Children with Reading Disability?, Scientific Studies of Reading*, *18*, 55–73. https://doi.org/10.1080/10888438.2013.836200

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*, *22*, 8–17.

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. https://doi.org/10.1016/j.jecp.2009.06.002

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, *16*(6), 504–525. https://doi.org/10.1080/10888438.2011.598199

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, *21*(5), 392–407. https://doi.org/10.1080/10888438.2017.1315424

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5–21. https://doi.org/10.1080/10888438.2013.819356

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing*, *7*(3), 295–326. https://doi.org/10.1007/BF03162082

Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A. M. (2012). From spelling pronunciation to lexical access: A second step in word decoding? *Scientific Studies of Reading*, *16*(4), 341–359. https://doi.org/10.1080/10888438.2011.568556

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* (pp. 169–190). John Benjamins. https://doi.org/10.1075/swll.15.10elb

Gibson, E. J. (1965). Learning to read: Experimental psychologists examine the process by which a fundamental intellectual skill is acquired. *Science*, *148*(3673), 1066–1072. https://doi.org/10.1126/science.148.3673.1066

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. https://doi.org/10.1207/S1532799XSSR0702_3

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. https://doi.org/10.1037/0033-295X.106.3.491

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. https://doi.org/10.1037/0033-295X.111.3.662

Hogan, T. P., Bowles, R. P., Catts, H. W., & Storkel, H. L. (2011). The influence of neighborhood density and word frequency on phoneme awareness in 2nd and 4th grades. *Journal of Communication Disorders*, *44*(1), 49–58. https://doi.org/10.1016/j.jcomdis.2010.07.002

Kearns, D. M. (2020). Does English have useful syllable division patterns? *Reading Research Quarterly*, *55*(S1), S145–S160. https://doi.org/10.1002/rrq.342

Kearns, D. M., Rogers, H. J., Koriakin, T., & Al Ghanem, R. (2016). Semantic and phonological ability to adjust recoding: A unique correlate of word reading skill? *Scientific Studies of Reading*, *20*(6), 455–470. https://doi.org/10.1080/10888438.2016.1217865

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). Erlbaum Publishers.

McCrimmon, A. W., & Smith, A. D. (2013). *Review of the Wechsler Abbreviated Scale of Intelligence, (WASI-II)*.

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual*. Woodcock-Johnson III.

Nation, K., & Castles, A. (2017). Putting the learning into orthographic learning. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 147–168). John Benjamins. https://doi.org/10.1075/swll.15.09nat

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. https://doi.org/10.1016/j.jecp.2009.05.001

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

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). Erlbaum.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbr, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). John Benjamins. https://doi.org/10.1075/swll.11.14per

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*(2), 31–74. https://doi.org/10.1111/1529-1006.00004

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, *36*(2), 321–339. https://doi.org/10.1016/0022-0965(83)90036-X

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

Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions teaching direct mapping of graphemes in texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading*, *22*(3), 225–247. https://doi.org/10.1080/10888438.2018.1427753

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, *23*(3), 383–404. https://doi.org/10.1016/S0022-5371(84)90270-6

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. *Advances in Child Development and Behavior*, *36*, 31–82. https://doi.org/10.1016/S0065-2407(08)00002-5

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.), *New directions in communication disorders research. Explaining individual differences in reading: Theory and evidence* (pp. 45–68). Psychology Press.

Siegelman, N., Kearns, D. M., & Rueckl, J. G. (2020). Using information-theoretic measures to characterize the structure of the writing system: The case of orthographic-phonological regularities in English. *Behavior Research Methods*, *52*(3), 1292–1312. https://doi.org/10.3758/s13428-019-01317-y

Stanovich, K. E. (1991). Word recognition: Changing perspectives. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. II, pp. 418–452). Erlbaum.

Steacy, L. M., Compton, D. L., Petscher, Y., Elliott, J. D., Smith, K., Rueckl, J. G., Sawi, O., Frost, S. J., & Pugh, K. R. (2019). Development and prediction of context-dependent vowel pronunciation in elementary readers. *Scientific Studies of Reading*, *23*(1), 49–63. https://doi.org/10.1080/10888438.2018.1466303

Steacy, L. M., Wade-Woolley, L., Rueckl, J. G., Pugh, K., Elliott, J. D., & Compton, D. L. (2019). 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. https://doi.org/10.1080/10888438.2019.1620749

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

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

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). 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 Studies of Reading*, *16*(2), 122–140. https://doi.org/10.1080/10888438.2010.542527

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

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*(1), 75–94. https://doi.org/10.1146/annurev-linguistics-030514-124832

Wagner, R. K., Torgesen, J. K., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing 2*. Pro-ed.

Wang, H.-C., Castles, A., & Nickels, L. (2012). Word regularity affects orthographic learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *65*, 856–864. https://doi.org/10.1080/17470218.2012.672996

Wang, H.-C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, *17*(5), 369–384. https://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. https://doi.org/10.3758/BF03197678

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence -Second Edition (WASI-II)*. 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*(3), e12577 https://doi.org/10.1111/desc.12577

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

Zipke, M. (2016). The importance of flexibility of pronunciation in learning to decode: A training study in set for variability. *First Language*, *36*(1), 71–86. https://doi.org/10.1177/0142723716639495

Received November 3, 2020
Revision received March 16, 2021
Accepted May 5, 2021 ■