








# Modeling Item-Level Spelling Variance in Adults: Providing Further Insights into Lexical Quality

Valeria M. Rigobon <sup>a,b,c,d</sup>, Nuria Gutiérrez <sup>c</sup>, Ashley A. Edwards <sup>c</sup>, Nancy Marenčin <sup>c,e</sup>,  
Matt Cooper Borkenhagen <sup>c,e</sup>, Laura M. Steacy <sup>c,e</sup>, and Donald L. Compton <sup>a,b,c,d</sup>

<sup>a</sup>Australian Centre for the Advancement of Literacy, North Sydney, NSW, Australia; <sup>b</sup>School of Education, Australian Catholic University, North Sydney, NSW, Australia; <sup>c</sup>Florida Center for Reading Research, Tallahassee, Florida, USA; <sup>d</sup>Department of Psychology, Florida State University, Tallahassee, Florida, USA; <sup>e</sup>School of Teacher Education, Florida State University, Tallahassee, Florida, USA

## ABSTRACT

**Purpose:** The lexical quality (LQ) hypothesis predicts that a skilled reader's lexicon will be inhabited by a range of low- to high-quality items, and the probability of representing a word with high quality varies as a function of person-level, word-level, and item-specific variables. These predictions were tested with spelling accuracy as a gauge of LQ.

**Method:** Item-response based crossed random effects models explored simultaneous contributions of person-level (e.g. participant's decoding skill), word-level (e.g. word's transparency rating), item-specific (e.g. participant's familiarity with specific word), and person-by-word interaction predictors (e.g. decoding by transparency rating interaction) to the spelling of 25 commonly misspelled irregular English words in 61 undergraduate university students ( $M = 19.4$  years, 70.49% female, 39.34% Hispanic, 81.97% White).

**Results:** Substantial variance among individuals in item-level spelling accuracy was accounted for by person-level decoding skill; item-specific familiarity, proportion of schwas correctly represented, and correctly identifying the word from its mispronunciation; and an interaction of transparency rating by general decoding skill.


**Conclusions:** Consistent with the LQ hypothesis, results suggest that one's ability to form a high-quality lexical representation of a given word depends on a complex combination of person-level abilities, word-level characteristics, item-specific experiences, and an interaction between person- and word-level influences.

## KEYWORDS

spelling; lexical quality; orthography; phonology; adults

Across individuals, words vary in the strength with which they are represented in the mind, which leads to variation in reading skill, including comprehension (C. Perfetti, 2007; C. A. Perfetti, 1985, 1992; C. Perfetti & Hart, 2001; C. A. Perfetti & Hart, 2002; C. Perfetti & Stafura, 2014). This idea is articulated by the Lexical Quality Hypothesis (C. Perfetti & Hart, 2001), which states, "Lexical quality (LQ) refers to the extent to which the reader's knowledge of a given word represents the word's form and meaning constituents and knowledge of word use that combines meaning with pragmatic features" (p. 359). A word's representation<sup>1</sup> is considered of high quality to the extent that it has a precise orthographic representation (i.e., spelling; see Andrews, 2012; Andrews & Treiman, 2015; Andrews et al., 2020; Hersch & Andrews, 2012; Perfetti, 1991; C. A. Perfetti, 1992), redundant phonological representations (see, A. Edwards et al., 2021; Elbro, 1998; Elbro & Jensen,

**CONTACT** Valeria M. Rigobon  [valeria.rigobon@acu.edu.au](mailto:valeria.rigobon@acu.edu.au)  Australian Centre for the Advancement of Literacy, 3-21 Berry Street, North Sydney, NSW 2060, Australia

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/10888438.2023.2234519>

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

2005; Goswami, 2000; C. Perfetti & Hart, 2001; Stafura & Perfetti, 2017), and meaning that is both specific and flexible (see Bolger et al., 2008; Braze et al., 2007; Hsiao & Nation, 2018; C. Perfetti, 2007). As such, high-quality lexical representations allow orthographic and phonological knowledge of a word to uniquely facilitate recognition and spelling of the word, respectively (see Castles et al., 2018). Evidence supports the notion that skilled reading relies heavily on high quality word representations that include well-specified orthographic, phonological, and semantic knowledge. C. A. Perfetti and Hart (2002) reported factor analysis results on 445 undergraduates who completed a battery of tasks that sampled orthographic, phonological, and semantic knowledge sources. Outcomes indicated that more skilled readers exhibited a 2-factor solution, one representing lexical form and the other meaning-comprehension, with word reading acting as a linking variable between the two factors. In less skilled readers, a 3-factor solution was favored with separate orthographic, phonological, and meaning-comprehension factors and pseudoword reading linking the orthographic and phonological factors. The differential factor structure that emerged across reading groups suggested a more integrated knowledge of orthographic and phonological structures for skilled readers than for less skilled readers. Results led Perfetti and Hart to conclude that, “the lack of integration of orthographic performance for low skilled readers suggests that spelling knowledge is not serving word reading in the same way it is for skilled readers” (C. A. Perfetti & Hart, 2002, p. 211). Since this original study, there have been numerous published experiments (using various methods) that support the legitimacy of the LQ hypothesis (for reviews of the evidence see Andrews & Treiman, 2015; Andrews et al., 2020; C. Perfetti, 2007; C. A. Perfetti, 2017; Van Dyke & Shankweiler, 2012); so much so, that LQ has become ubiquitous within the reading literature.

The LQ hypothesis stipulates that a reader’s lexicon will be inhabited by a range of low- to high-quality items, with the mix of item quality varying across individuals and words (Perfetti, 1991, C. A. Perfetti, 1992). High-quality items are not widely permissive of influences beyond exceedingly constrained sources internal to the lexicon (i.e., word-specific spelling, pronunciation, and meaning) in skilled word recognition. Lower-quality items, on the other hand, are less stable and lack specified representations of a word’s orthography, phonology, and semantics, leading to incorrect spelling, delayed word recognition, and lower comprehension of text. C. A. Perfetti and Hart (2002) illustrate the variability of lower quality representations with this example (p. 192):

Presented with the word *incarcerate*, the reader. . .

- (a) pronounces it accurately and knows it has some negative meaning, but is not sure what that meaning is.
- (b) stumbles on its pronunciation, producing something like “in-cark-rate.”
- (c) pronounces it accurately and indicates that it means something like “to confine in prison,” but when attempting to speak the word to produce a message about someone going to jail, sometimes produces “incarcerate” and sometimes something more like “incarsate.”
- (d) can perform all the tasks in a, b, and c above but can spell the word correctly only on some attempts.

Perfetti and Hart stress that this final case is of particular importance to most individuals of high literacy skill – “the feeling of semantic and phonological competence coupled with a spelling block” (C. A. Perfetti & Hart, 2002, p. 192). Thus, C. A. Perfetti (1992) and colleagues (C. Perfetti & Hart, 2001; C. A. Perfetti & Hart, 2002) suggest that the key measure for gauging LQ across the continuum of literacy skill is spelling. Specifically, accurate spelling requires recall of a high level of detailed knowledge about the lexical representation while many reading tasks can be achieved with recognition based on partial lexical information supplemented by context (see Bosman & Van Orden, 1997; Ehri, 1997). To test important tenets of the LQ hypothesis, the current study explores the roles of theoretically driven person-level, word-level, item-specific, and person-by-word interaction predictors of an

individual's likelihood of spelling a complex English word correctly, which serves as a proxy for an individual's item-level lexical quality.

Perfetti (1991, C. A. Perfetti, 1992) proposes that one of the critical events for an item's transition across the continuum from low to high quality is for a representation to become fully specified in spelling, pronunciation, and meaning. In spelling, this means a given phonological representation (i.e., spoken form of a word) will be sufficient to activate a specific spelling with its precise orthographic units in the correct order, perhaps with more effort than word recognition in reading (Burt & Fury, 2000). Andrews and colleagues have provided compelling evidence supporting Perfetti's claim that orthographic precision, as referenced by spelling ability, is a critical index of a word's LQ within an individual. For instance, in a recent study of 785 undergraduate students, Andrews et al.'s (2020) factor analyses and principal component analysis demonstrated that spelling skill consistently represents a separate lexical component that is partially independent of variance in reading comprehension, reading speed, and vocabulary, concluding that this component aligns with the precision dimension of LQ.

The other critical element of high-quality representations is phonological redundancy (see Perfetti, 1991, C. A. Perfetti, 1992; C. A. Perfetti & Hart, 2002). Redundancy refers to multiple or "redundant" stored phonological representations associated with a single high-quality lexical representation; comprised of "one from spoken language and one recoverable from orthographic-to-phonological mappings" of the word (p. 190, C. A. Perfetti & Hart, 2002). This assumes that decoding an unknown letter string during orthographic learning (see Elbro & de Jong, 2017; Elbro et al., 2012; Nation & Castles, 2017; Share, 2008) results in the representation of the word's decoded form (i.e., "spelling pronunciation" [Elbro et al., 2012]; "overpronunciation" [Holmes & Malone, 2004; Ormrod & Jenkins, 1989]; "regularized pronunciation" [T. Ocal & L. C. Ehri, 2017b]) which influences the overall quality of the representation over time. For example, in addition to the spoken form of "tongue,"/ton-goo /may be another pronunciation derived from English decoding rules that is activated when attempting to read/spell "tongue" and allows for the irregular word "tongue" to be recognized and spelled accurately. Phonological redundancy can likely take other forms as well (e.g., overlap of orthographic and phonological forms of cognates across languages sharing a common alphabet; see Rigobon et al., 2023). Precise and redundant representations yield lexical retrieval that is reliable and coherent in the sense that the orthographic, phonological, and semantic knowledge of a word are "available synchronously at retrieval, giving the impression of a unitary word perception event" (C. Perfetti & Hart, 2001, p. 69).

One way to test the availability of a redundant phonological representation in the lexicon is to measure a person's **set for variability (SfV)**, which has been conceptualized as the process of disambiguating the mismatch that can occur between a familiar word's phonological representation and the decoded form of the word (see Elbro et al., 2012), especially for the purpose of arriving at the word's correct pronunciation during word reading (Gibson & Levin, 1975; Venezky, 1999; Tunmer & Chapman, 2012). To test this process, participants are orally presented with known English words that are "mispronounced" based on standard decoding rules (e.g.,/breek-fast/for "breakfast"), and asked to provide the correct pronunciation aloud. Total SfV performance has been reported as a significant predictor of school-age children's regular word reading in Dutch (Elbro et al., 2012) and general word reading skill in English (Kearns et al., 2016; Steacy et al., 2022; Tunmer & Chapman, 2012), as well as nonword reading in English (Steacy, Compton, et al., 2019). Further, item-specific SfV has been shown to be significantly predictive of variance in grade 2–5 students' irregular word reading (Steacy, Wade Woolley, et al., 2019). If spelling accuracy is the key measure for gauging LQ and high LQ is dependent on both precision and redundancy, then SfV performance should be predictive of spelling accuracy at the level of the item.

The development of high-quality representations is a gradual process that operates at the item level (Nation & Castles, 2017; Share, 1995), with LQ differing across words for a given individual and across individuals for a given word (C. A. Perfetti & Hart, 2002). At the person level, the LQ hypothesis acknowledges that skilled readers have an advantage over less skilled readers in their ability to add new

information about spelling, pronunciation, or meaning of an impoverished representation via superior foundational resources (e.g., decoding, spelling, grammatical, and vocabulary skills). The current study tests this assumption by including general SfV performance, English decoding ability, and English word familiarity as person-level predictors of spelling accuracy. These allow us to explore whether important person-level foundational resources associated with the general availability of redundant phonological representations, decoding skill (i.e., proxy for general knowledge of orthographic-phonological relationships), and general semantic knowledge of words will distinguish those with precise orthographic representations, as gauged by item-level spelling performance.

The LQ hypothesis acknowledges that even skilled readers will have low-quality representations for many words, both low frequency words from general vocabulary and words from specialized vocabularies (see Andrews, 2012; Andrews & Treiman, 2015; C. Perfetti & Hart, 2001; C. A. Perfetti & Hart, 2002). In a quasiregular orthography like English, word characteristics such as regularity can also impact LQ, given that the “mismatch distance” between a word’s decoded form and its phonological representation in a reader’s lexicon can lead to more difficulty in accurate reading or spelling of irregular words compared to regular words (see Nation & Snowling, 1998). As such, frequency and transparency of spelling to pronunciation are included as general word-level predictors of spelling accuracy in the current study.

While the LQ hypothesis was developed to explain variations in reading ability, it should not be equated with deficit hypotheses that look to explain individual differences in reading skill by variations in component skills such as phonological processing, naming skill, and semantic knowledge (C. Perfetti, 2007). Instead, readers build representations from repeated encounters with a word, leading to item-specific knowledge across orthography, phonology, and semantics. Therefore, the current study explores the effects of the following item-specific predictors: SfV (i.e., if the participant was able to offer the real pronunciation of a target word given its mispronunciation), the proportion of correctly spelled schwas in the word, and familiarity with the spoken form of a specific word. These predictors allowed us to test the hypothesis that precise orthographic representations are associated with three important connections between orthography, phonology, and semantics. Item-specific SfV captures the presence of a recoverable decoded form derived from the orthographic-to-phonological mappings of the word, which is presumed to be a precursor to the formation of a high-quality representation. The proportion of correctly spelled schwas in the word captures specific and precise connections between phonology and orthography that are ambiguous and unrecoverable using frequent orthographic-phonological relations. While item-specific familiarity may not capture deep semantic knowledge of a word, it does capture information about how an individual represents the link between phonological and semantic knowledge of a word in the lexicon.

Finally, highly skilled readers are more likely to fully decode unfamiliar or unknown words in text (Perfetti, 1991, 1922), leading to the successful formation of more high-quality lexical entries and continuous “tuning” (see Andrews, 2012; Castles et al., 2007; Hersch & Andrews, 2012) of those representations with more reading and writing experience. As skilled readers continue to update and tune representations, their knowledge about orthographic-phonological relationships—based on observation of co-occurring patterns and regularities of the English orthography and nuances of a word’s meaning in different contexts—continues to evolve, with this tuning resulting in the addition of more high-quality representations to the lexicon. Consequently, skilled readers with better decoding skills should have a lower reliance on transparent orthographic-phonological relationships for accurate spelling and benefit from a richer set of lexical resources when trying to supplement a lower quality representation during reading and spelling. Adults with poorer decoding skills, on the other hand, should be less likely to fully decode words in text and therefore, store fewer high-quality representations in the lexicon. With fewer high-quality representations to draw from when reading and spelling, poor decoders should then be more likely to exhibit a higher reliance on more transparent and frequent orthographic-phonological relationships that leads to more partial decoding, along with effortful and inaccurate spelling. To test these assumptions about the bidirectional nature of this relationship between decoding skill and addition of high-quality lexical representations, the

current study explores the contributions of a person-by-word interaction to spelling accuracy that would be consistent with the formation of precise orthographic representations within the LQ hypothesis: the interaction between decoding skill and word spelling-to-pronunciation transparency rating.

Overall, previous work from Andrews and colleagues has demonstrated that spelling accuracy is a useful gauge of precision in quality of lexical representations, while others (e.g., Burt & Fury, 2000; Holmes & Malone, 2004; T. Ocal & Ehri, 2017a; T. Ocal & L. C. Ehri, 2017b; Ormrod & Jenkins, 1989) have explained general spelling performance in adults as a function of either general person-level skills or general word characteristics. However, the use of item-level modeling to capture individual differences in item-level development of high-quality representations has yet to be applied to spelling performance. Additionally, the role of **redundancy** in quality of lexical representations has yet to be empirically tested with consideration of both person- and word-level influences. Thus, the current study extends previous work by modeling predictors of individual differences in adults' item-level spelling performance (i.e., accuracy of spelling responses to individual words instead of total accuracy on a spelling task) as a gauge of LQ<sup>2008</sup>, which affords consideration of simultaneous person-level, word-level, item-specific, and person-by-word-level influences on LQ.

## Present Study

We asked undergraduate adult participants to spell a subset of commonly misspelled English words (from the Macquarie University Advanced Adult Spelling Test, Caruana et al., 2019) with complex phonological to orthographic relationships (e.g., *bureaucracy*) and subsequently explored the role of important person-level, word-level, and item-specific predictors in forecasting individual differences in item-level spelling performance. Words were intentionally chosen so that spelling accuracy could not be achieved by only applying regular orthographic-phonological correspondences, thereby making spelling accuracy reliant on recall of stored lexical representations. We purposely selected predictors that allowed us to assess certain assumptions inherent to the concepts of orthographic precision and phonological redundancy as they relate to LQ. In doing so, we are mindful of C. A. Perfetti's (2017) warning that LQ is a theoretical framework rather than a theory, expressing general claims based on key constructs that can be elaborated into testable theories that Perfetti refers to as a "hypothesis-testing agenda" (C. A. Perfetti, 2017, p. 53). In our hypothesis-testing process, person-level variables were selected to help identify important component skills that distinguish those who have greater skill in establishing high- versus low-quality representations. Word-level predictors were selected to account for characteristics of the target words that might make them more likely to be represented with high quality. Item-specific predictors were selected to capture connections between orthographic, phonological, and semantic knowledge of a specific individual's representation of a given word. Our approach also allowed us to explore a key person-by-word interaction of interest that would be consistent with the formation of precise orthographic representations within the LQ hypothesis: the interaction between decoding skill and word spelling-to-pronunciation transparency rating.

In sum, the LQ hypothesis predicts that a skilled reader's lexicon will be inhabited by a range of low- to high-quality items, and further that the probability a representation will be high-quality varies as a function of person- and word-level characteristics. In addition, because the development of high-quality representations is a gradual tuning process that operates at the item level, the LQ hypothesis postulates that including item-specific predictors is essential for modeling a reader's representational knowledge. Thus, we test the following predictions regarding spelling performance based on the tenets of the LQ hypothesis: 1) large variation will exist across person and word regarding which items are considered to have high representational quality (as exhibited by the number and diversity of spelling errors across words and persons); 2) measures representing an individual's foundational resources related to forming high-quality representations (e.g., decoding skill) should make a contribution to the probability that a word is spelled correctly above and beyond item-specific predictors; 3) word

characteristics that affect the likelihood of encountering a word (e.g., frequency) and ease of successfully decoding it (e.g., spelling-to-pronunciation transparency) should contribute to the probability that a word is spelled correctly; 4) item-specific predictors representing links between orthographic, phonological, and semantic knowledge of a word will contribute significantly to individual differences in the probability of spelling accuracy; and 5) poor decoders will be more dependent on transparent orthographic-phonological relationships in a word to generate a correct spelling.

## Methods

### Participants

Due to the SARS-CoV-2 pandemic, data was collected remotely using Zoom, Microsoft PowerPoint, and Qualtrics services from 61 undergraduate students (ages 18–24 years) in the psychology subject pool of a large public university in the Southeast region of the U.S. Prior to the study's initiation, ethical approval was obtained from the Florida State University's ethics committee for human subject research, in compliance with the U.S. Federal Policy for the Protection of Human Subjects. Individual consent was obtained online for each participant before the Zoom testing session began. Participating subjects who identified as proficient English speakers completed one remote Zoom testing session of no more than 90 minutes, and they were compensated with extra credit for selected courses. Demographic data for participants are presented in [Table 1](#).

### Procedures

Each participant was remotely assessed on their familiarity with and spelling of 25 words from the Macquarie University Advanced Adult Spelling Test (MAAST; Caruana et al., 2019) list of commonly misspelled English words, which varied in length, print frequency, number of morphemes and schwas, and transparency of spelling to pronunciation. General word reading and decoding skills were assessed

**Table 1.** Demographic characteristics of participants.

Variable	Full sample <i>N</i> = 61
Age (Years)	<i>M</i> = 19.4 <i>SD</i> = 1.55
Subjective Spelling Ability Rating	<i>M</i> = 2.59 <i>SD</i> = 0.72
Sex (%)	
Female	70.49%
Male	29.51%
Ethnicity (%)	
Hispanic/Latino	39.34%
Non-Hispanic/Latino	60.66%
Multilingual (%)	
Yes	37.70%
No	62.30%
Primarily English-speaking (%)	
Yes	90.16%
No	9.84%
Race (%)	
American Indian/Alaskan Native	0.00%
Asian	3.28%
Black/African American	8.19%
White	81.97%
Multiracial	6.56%

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. Subjective spelling ability rating ranged on a scale from 1 to 4, with 1 being a very bad speller and 4 being a very good speller.

using online materials adapted from standardized tests. The set for variability and English familiarity tests were adapted to be administered online via Qualtrics surveys from measures used by Steacy, Compton, et al. (2019) and Kearns et al. (2016), respectively. Raw scores were used for each measure for descriptive information and data analysis, with scores grand mean-centered for the statistical models reported in the results.

Research assistants received weekly training and practice sessions on Zoom for two months, after which each assistant completed mock test administration sessions online with the trainer, who addressed any incorrect/incomplete testing behaviors after the session was conducted. A fidelity-of-implementation checklist was developed based on the testing scripts for the standardized assessments and the researcher-created measures. The mean fidelity estimate was 98% across testers in their mock test administration sessions prior to testing participants. Administration and scoring procedures for the testing sessions were periodically checked to ensure that fidelity was maintained at > 90% across participants over the course of data collection. The REDCap (Research Electronic Data Capture) tool hosted at Vanderbilt University (Harris et al., 2009) was used to enter and manage the data throughout the study period.

## Measures

In the present study, we make a distinction between person-level, word-level, and item-specific predictors used in the analytic models. **Person-level** measures refer to a participant's total score on a given task (e.g., decoding performance) instead of a single response to a single item; **word-level** measures refer to those that remain *fixed* across participants' individual responses, as these are general characteristics of the words being presented to the participants (e.g., frequency, number of morphemes); and **item-specific** measures are those that vary across participants' individual responses to a single word within a task. For example, an item-specific predictor is a participant's self-reported familiarity with a specific word (e.g., *etiquette*) in the target spelling task (0 = unfamiliar, 1 = familiar with the pronunciation of *etiquette*); the number of morphemes in *etiquette* is coded at the word level as a 1 across all participants' responses; and lastly, the participant's total score on the familiarity task is coded at the person level as a number between 0 and 25 based on how many words were identified as being familiar to the individual participant.

### Dependent measure

**Target spelling.** To measure participants' ability to spell complex words, 25 words from the Macquarie University Advanced Adult Spelling Test (MAAST; Caruana et al., 2019) list of commonly misspelled English words were chosen for containing at least one *schwa* (unstressed vowel within a word, /ə/). Participants were guided to disable their spell check before beginning the task. Following a traditional dictation method, participants heard via recording: each target word read aloud, the word used in a sentence, a repetition of the word, and instruction to type their response in the corresponding space on a survey. Each spelling response was scored as 0 for incorrect (i.e., any deviation from the correct spelling in American English) or 1 for correct (i.e., all letters provided in the correct order with no missing or additional letters, according to American English spelling). The dichotomous score for each participant's spelling of a single word was predicted at the item level in the models by the following item-specific, person-level, and word-level variables.

### Item-specific measures

**Proportion of correctly spelled schwas.** The target spelling words included 13 words with 1 schwa and 12 words with 2 schwas based on the English Lexicon Project (ELP; Balota et al., 2007).<sup>3</sup> The proportion of schwas correctly spelled was calculated by dividing the number of schwas correctly spelled by a single participant in a given word (e.g., *soliloquy*) by the number of total schwas in the word's correct spelling (e.g., 0 for misspelling as *syllilaqui*, 0.5 for misspelling as *syloiquy*, 1 for *soliloquy*, *soliloqui*). This was included as an item-specific predictor in the final models to gauge an

individual's representation of a specific and precise phonological to orthographic correspondence which cannot be easily recovered by relying on frequent orthographic-phonological relationships.

**Set for Variability (SfV).** Based on the work of Tunmer and Chapman (1998, 2012) and Steacy, Compton, et al. (2019) with elementary-aged students, SfV was evaluated by participants' ability to derive the correct pronunciation from spoken English words that are "mispronounced" based on regular decoding rules, as they might be if they were regular words or partially decoded (e.g., /sɛpəɹɪt/ for /sɛpɪət/). This is an experimental measure aimed at capturing an individual's ability to access a redundant orthographic-to-phonological mapping (i.e., decoded form) for each target spelling word. Responses were coded dichotomously; a score of 1 was assigned for the correct pronunciation of a target item (i.e., "separate" in adjective form) and a score of 0 for a response of any other English word, a nonword, or "I don't know".

**Target word familiarity.** This measure was adapted from a measure of polymorphemic words (Kearns et al., 2016) to the list of 25 target spelling words and accounted for individual differences in prior exposures to the target spelling words, several of which have low frequency of appearing in text. In the dependent spelling task instructions, participants were asked to pause after attempting to spell a word and provide a response for whether that word sounded familiar or not (based on having heard or encountered the word in spoken conversation or written text prior to the testing sessions). Responses were coded dichotomously (1 = "yes," 0 = "no" or "I'm not sure.").

#### **Person-level measures**

**Set for Variability (SfV) total.** This score represented an individual's total correct responses to the 25 target spelling words and 20 words used in prior administrations of the SfV task (Kearns et al., 2016; Steacy, Compton, et al., 2019; Tunmer & Chapman, 1998, 2012) for a total of 45 items. Cronbach's alpha was .80.

**Target word familiarity.** We consider an individual's total score on this measure to be a proxy for semantic knowledge of the 25 target spelling words. Cronbach's alpha was .86.

**Decoding fluency.** Participants were asked to read a list of 66 nonwords in English as quickly and accurately as possible within the span of 45s in the Phonemic Decoding Efficiency subtest from the TOWRE-2 (Torgesen et al., 2012). The authors report an alternate forms reliability of .92.

#### **Word-level measures**

**Frequency.** Target spelling words' log-transformed HAL frequency values, based on the Hyperspace Analogue to Language corpus, were taken from the ELP (Balota et al., 2007). The log transformed HAL frequency reported for the list of target words ranges from 3.219 to 10.498.

**Number of morphemes.** This measures the number of morphemes in each word based on hand coding completed by an experienced speech language pathologist. The range reported for the target spelling words is 1 to 3 morphemes.

**Spelling-to-Pronunciation Transparency (STPT) rating.** To address the distance between an irregular word's correct pronunciation and its decoded form (i.e., how easily a word's correct pronunciation can be derived from traditional orthographic-phonological relationships based on a word's spelling), we used word ratings from A. A. Edwards et al. (2021) database. Raters were instructed to pretend that a letter string was unfamiliar to them, try applying a letter-to-sound reading strategy to the letter string, and rate the difficulty of matching a word's decoded form to the word's standard pronunciation (provided via audio recording for reference) on a Likert scale from 1 to 6 (1 = very easy to match, 6 = very difficult to match). Ratings from



this database, including ratings for the 25 target spelling words in this study, have been shown to be reliable based on their high correlation with expert ratings reported by Steacy et al. (2017). The range of ratings reported for the target spelling words is 2.11 to 5.30.

### Data analytic procedures

Item-response based crossed random effects models were used to simultaneously account for the role of person-level, word-level, item-specific, and person-by-word interaction predictors of item-level word spelling variance. The models are “cross-classified” because responses to the spelling task originate from the intersection of each participant responding to each word in the same set of spelling items, where words are treated as a random factor that allows variability in spelling to be explained by word-level characteristics (e.g., frequency) and participants are treated as a random factor that allows variability in spelling accuracy to be explained by person-level characteristics (e.g., decoding skill). Simulation studies from Cho et al. (2012) have demonstrated the robustness of these models to sample size and number of items, “resulting in more power than typical individual regression models” (Steacy, 2020, p. 156).

These cross-classified models were used to predict the participants’ spelling of a specific item (e.g., *medieval*) coded as a dichotomous response (1=correct, 0=incorrect) using person-level characteristics (e.g., decoding skill), general word-level characteristics (e.g., transparency rating of *medieval*), item-specific (i.e., person-by-word) responses to other tasks (e.g., familiarity with spoken form of *medieval*), and a person-by-word level interaction (e.g., phonemic decoding efficiency by word transparency rating) as predictors. These models assume nonindependence in the data due to participants’ responses to the same set of words in the target spelling measure (i.e., nested responses within both person and word), which is accounted for by the inclusion of a by-subject random intercept and by-item random intercept in the models (see Brauer and Curtin, 2018). All reported analyses were conducted using a binomial distribution with a logit link, available through the `glmer` function in the `lme4` package (Bates et al., 2015) from R programming (R Development Team, 2012). All continuous person and word predictors were grand mean-centered to aid in interpretation of the intercept and coefficients. The R code used to conduct the item-response based crossed random effects models can be found on Open Science Framework through the link provided in the supplemental materials.

## Results

Descriptive statistics are provided in Table 2 for total target word spelling, person-level, and word-level variables. The mean number of words correctly spelled across individuals was just under 9, suggesting that, overall, the target spelling items were challenging. The variance in spelling

**Table 2.** Descriptive statistics and zero-order correlations of person & word-level features in the full sample (N = 61).

Person-Level Variable	<i>M</i>	<i>SD</i>	1	2	3	4
1. Decoding (PDE)	49.41	9.29				
2. General Set for Variability	35.46	5.62	.50**			
3. Target Familiarity (Total)	18.86	3.52	.44**	.66**		
4. Target Spelling (Total)	8.89	4.95	.52**	.55**	.48**	
Word-Level Variable						
1. Length	9.44	1.30				
2. Log HAL Frequency	6.32	1.67	.05			
3. Number of Morphemes	2.00	0.85	.44*	-.17		
4. Number of Schwas	1.48	0.50	.29	-.19	.47*	
5. STPT Rating	3.74	0.80	.28	-.47*	.18	.14

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. \* indicates  $p < .05$ . \*\* indicates  $p < .001$ . PDE = Phonemic Decoding Efficiency subtest from the TOWRE-2 (Torgesen et al., 2012); HAL = Hyperspace Analogue to Language corpus; STPT = Spelling to Pronunciation Transparency.

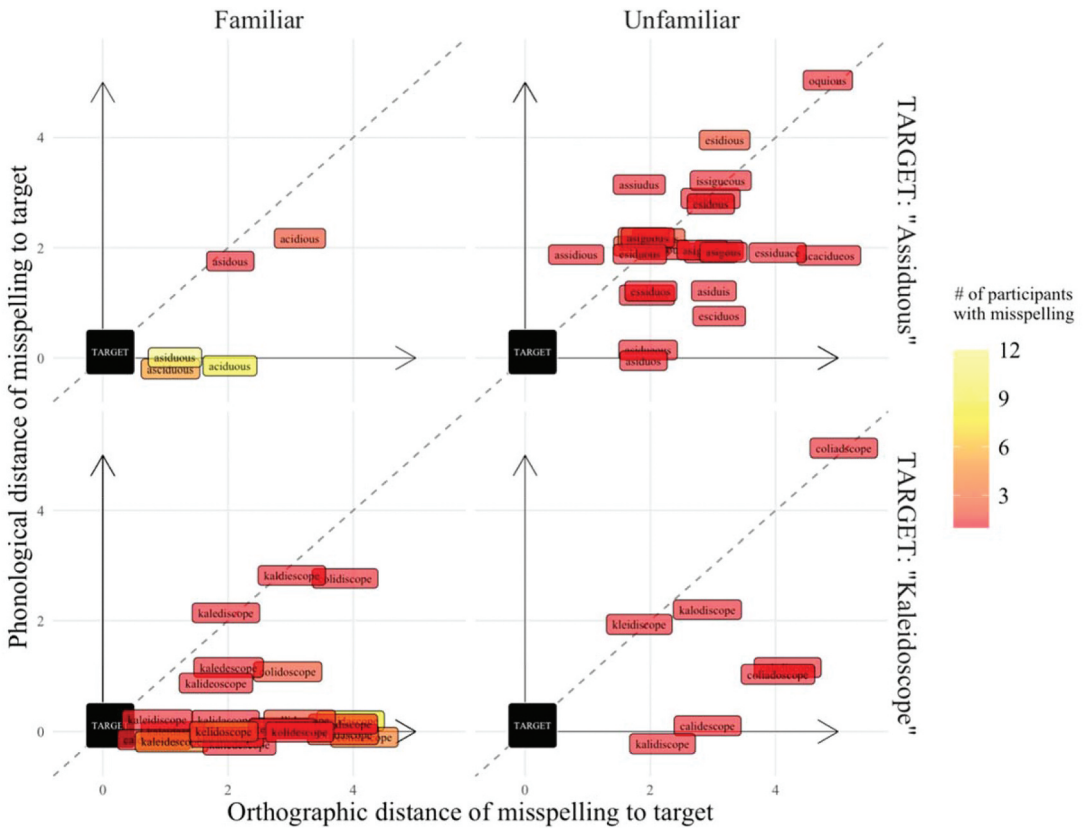
**Table 3.** Difficulty ranking by number of misspellings per target spelling word (N = 25).

Word	Difficulty Ranking	N of Participants Who Misspelled Word	N of Unique Misspellings
assiduous	1	59	30
bureaucracy	2	56	32
soliloquy	3	52	40
kaleidoscope	4	52	30
charlatan	5	51	28
effeminate	6	49	23
cacophony	7	46	33
debauchery	8	46	27
aqueduct	9	46	9
colloquial	10	44	18
plagiarism	11	44	12
embarrassed	12	43	9
omniscient	13	42	16
medieval	14	40	14
nauseous	15	38	29
miscellaneous	16	38	25
boisterous	17	36	22
poignant	18	35	22
pneumonia	19	34	21
separate	20	30	2
etiquette	21	26	26
pinnacle	22	24	13
nonchalant	23	18	14
avalanche	24	10	5
tranquil	25	8	6

Note. The maximum number of individuals who could misspell a given word was 61, the total sample size.

performance was noteworthy, with a standard deviation of just under 5 words. These results prompted a closer look at the overall difficulty of each word and the number of unique misspellings for each word (see Table 3), where substantial differences in the difficulty of correctly spelling a word are highlighted in *soliloquy* (spelled correctly by only 9 of 61 participants) compared to *avalanche* (spelled correctly by 51 participants). In addition, there was wide variation in the number of unique incorrect spellings by participants, with *soliloquy* having 40 unique misspellings whereas *avalanche* only 5. Interestingly, the number of unique misspellings per word correlated strongly with the transparency ratings ( $r = .70$ ,  $p < .001$ ), suggesting that the varying degrees of ambiguity in these target words' spelling-to-pronunciation transparency may better represent an important source of spelling difficulty than other word features that are traditionally associated with word reading and spelling difficulty, such as length and frequency, which were only moderately correlated with the number of unique misspellings at  $r = .27$  and  $r = .43$ , respectively. To help conceptualize and explain the sheer number of unique spelling errors across words we provide a visualization of the unique errors for the most frequently misspelled word *assiduous*, compared to *kaleidoscope*, the fourth most difficult word to spell, each of which has 30 unique misspellings (see Figure 1). The figure depicts misspellings of each word across three important dimensions: orthographic and phonological distance from the correct spelling (measured using Damerau-Levenshtein distances) and each participant's item-specific familiarity with the target word.

Zero-order correlations for person and word features are provided in Table 2. At the word level, number of morphemes showed a moderately strong positive correlation with number of schwas, indicating that more morphologically complex words tend to have a higher number of schwas in the target spelling words. Word frequency, on the other hand, showed a moderately strong negative correlation with subjective ratings of spelling-to-pronunciation transparency, which suggests that words with higher transparency tend to appear more frequently in text. All other word-level correlations were insignificant. At the person level, moderately strong positive correlations are reported for total scores on target familiarity and word identification; familiarity and decoding; and SfV and



**Figure 1.** Visualization of misspellings for “assiduous” and “kaleidoscope” on target spelling task. *Note.* Here the points (labels) are jittered to minimize overlapping. The jittering introduces randomness into rendering the plot. The reason for this jittering is that without it the misspelled forms would overlap at each level of x- and y- such that you could only see one misspelling, and the rest would be plotted behind it. The left panels show misspellings where the participant rated the target as familiar, and the right panels show those corresponding to unfamiliar ratings for the target. Target words were chosen because they exhibited variability in the familiarity rating across participants, where “assiduous” was only familiar to 11 participants and “kaleidoscope” was reported as being familiar to 48 participants. The origin of each plot represents the target itself, such that the distance of a given misspelling from the origin can be interpreted as the extent to which that misspelling is dissimilar from the target word either with respect to its orthographic structure (x-axis) or phonological structure (y-axis). Word labels (points) are colored based on the number of particular misspellings observed across participants. Dashed diagonals are included as a reference.

decoding. Total scores on target spelling also had a moderately strong positive correlation with total familiarity, decoding, and SfV. The strongest positive correlation is reported for total scores on familiarity and SfV, which was expected given that the two experimental tasks contain overlapping items.

A series of models (i.e., Unconditional, Main Effects I and II) are presented in Table 4, illustrating separate sources of variance in item-specific word spelling accuracy associated with item-specific, person-level, and word-level predictors. Probabilities of correct spelling were calculated based on the logit estimates from each respective model. The Unconditional Model’s intercept ( $\gamma = -0.85$ ,  $z = -2.75$ ,  $p = .01$ ) indicated that the average probability of a correct response across words and participants on the target spelling task was 29.95%. Variance estimates at the person level (1.372) and word level (1.692) suggested that there was significant variance to be explained at both levels in subsequent models. Next, all person- and word-level predictors were entered into Main Effects Model I simultaneously to predict spelling performance. Only SfV and decoding had significant main effects at the person level, and no predictors at the word level significantly accounted for variance in likelihood of spelling accuracy. Item-specific measures of familiarity, SfV, and proportion of correctly

**Table 4.** Fixed effects and variance estimates predicting probability of correct word spelling responses on target spelling task.

Fixed effects	Unconditional Model				Main Effects Model I				Main Effects Model II			
	Est.	SE	z	p	Est.	SE	z	p	Est.	SE	z	p
<b>Intercept</b>	<b>-.85</b>	<b>.31</b>	<b>-2.75</b>	<b>.01</b>	<b>-.85</b>	<b>.25</b>	<b>-3.36</b>	<b>&lt;.001</b>	<b>-12.83</b>	<b>1.44</b>	<b>-8.89</b>	<b>&lt;.001</b>
<b>Item-Specific Factors<sup>a</sup></b>												
Proportion of CSS	–	–	–	–	–	–	–	–	<b>11.70</b>	<b>1.40</b>	<b>8.37</b>	<b>&lt;.001</b>
SfV	–	–	–	–	–	–	–	–	<b>.86</b>	<b>.29</b>	<b>2.92</b>	<b>&lt;.01</b>
Target Familiarity	–	–	–	–	–	–	–	–	<b>1.07</b>	<b>.34</b>	<b>3.18</b>	<b>&lt;.01</b>
<b>Person-Level Factors<sup>b</sup></b>												
Decoding (PDE)	–	–	–	–	<b>.04</b>	<b>.02</b>	<b>2.75</b>	<b>.01</b>	<b>.05</b>	<b>.02</b>	<b>3.17</b>	<b>&lt;.01</b>
SfV	–	–	–	–	<b>.07</b>	<b>.03</b>	<b>2.07</b>	<b>.04</b>	.02	.03	.59	.56
Target Familiarity	–	–	–	–	.05	.05	1.09	.28	.05	.05	.98	.33
<b>Word-Level Factors<sup>c</sup></b>												
Log HAL Frequency	–	–	–	–	.13	.16	.82	.41	–.21	.20	–1.03	.30
Morphemes	–	–	–	–	–.48	.28	–1.73	.08	–.54	.35	–1.52	.13
STPT Rating	–	–	–	–	–.44	.32	–1.35	.18	–.74	.41	–1.81	.07
Intercepts		Variance				Variance						
Person		1.37				0.63						
Word		1.69				1.19						
							Variance Explained					
							49.15%					
							29.43%					

Note. Est. = Parameter Estimate; SE = Standard Error; PDE = Phonemic Decoding Efficiency subtest from the TOWRE-2 (Torgesen et al., 2012); SfV = Set For Variability; CSS = Correctly Spelled Schwas; HAL = Hyperspace Analogue to Language corpus; STPT = Spelling to Pronunciation Transparency. Each of the predictors and respective estimates represent the results from predicting probability of word spelling accuracy from all variables simultaneously (i.e., in the presence of all other word- and person-level predictors in the model). <sup>a</sup>Item-specific factors represent item-specific performance; <sup>b</sup>Person-level factors represent aggregate performance by the individual on the measure. <sup>c</sup>Word-level factors represent fixed characteristics of the 25 target spelling words that remain constant across individual participants.

spelled schwas were added to the model (see Main Effects Model II). In this model all three item-specific predictors were significant ( $p < .01$ ), indicating that an individual was more likely to spell a word correctly if they a) were familiar with the word's spoken form prior to the testing session, b) were able to correctly identify the word from hearing its mispronounced form, or c) spelled the letters representing the schwa(s) within the word correctly. Similar to Model I, person-level decoding skill was a significant predictor, and there were no significant word-level predictors of item spelling accuracy in Main Effects Model II. Note that while item-specific predictors of SfV and familiarity are competing for variance with person-level measures of SfV and familiarity in Main Effects Model II and the following exploratory models, the point biserial correlation between a participant's response to one specific item on the SfV task and their total score on the SfV task varies greatly depending on the item (i.e., lowest correlation is .07 for a participant's total SfV score and response to the word *miscellaneous* and the highest correlation is .65 for *omniscient*). This applies to the familiarity measure as well, with point biserial correlations ranging from .11 for *pneumonia* to .69 for *soliloquy*. These results demonstrate that none of the single items can explain all the variance in overall performance, suggesting that item-specific and overall scores provide unique information that are not totally overlapping. Therefore, the predictive power of item-specific SfV or familiarity responses does not depend entirely on general performance for these measures, nor does the predictive power of general performance on these measures depend entirely on responses to specific items in SfV or familiarity measures.

Next, two exploratory models were conducted to explore the person-by-word interaction between decoding skill and word transparency rating (see Table 5). We first provide results from Exploratory Model I in which we added the interaction of decoding skill by word transparency rating to Main Effects Model II. The interaction in Exploratory Model I was not a significant predictor of spelling ability. However, given that the proportion of correctly spelled schwas in the target spelling words is necessary (but not sufficient) for accurate spelling in the outcome variable, we must acknowledge a dependency between this predictor and the outcome variable in Main Effects Model I and Exploratory Model I. We considered this a critical predictor to include in the models because it

**Table 5.** Exploratory interaction effects predicting probability of correct word spelling responses on target spelling task.

Fixed effects	Unconditional model				Exploratory Model I				Exploratory Model II			
	Est.	SE	z	p.	Est.	SE	z	p	Est.	SE	z	p
<b>Intercept</b>	<b>-.85</b>	<b>.31</b>	<b>-2.75</b>	<b>.01</b>	<b>-12.66</b>	<b>1.43</b>	<b>-8.85</b>	<b>&lt;.001</b>	<b>-2.59</b>	<b>.35</b>	<b>-7.46</b>	<b>&lt;.001</b>
<b>Interaction<sup>a</sup></b>												
PDE x STPT Rating	–	–	–	–	.01	.01	1.18	.24	<b>.03</b>	<b>.01</b>	<b>2.76</b>	<b>.01</b>
<b>Item-Specific Factors<sup>b</sup></b>	–	–	–	–					–	–	–	–
Proportion of CSS	–	–	–	–	<b>11.52</b>	<b>1.38</b>	<b>8.33</b>	<b>&lt;.001</b>	–	–	–	–
SfV	–	–	–	–	<b>.84</b>	<b>.29</b>	<b>2.86</b>	<b>&lt;.01</b>	<b>.75</b>	<b>.24</b>	<b>3.17</b>	<b>&lt;.01</b>
Target Familiarity	–	–	–	–	<b>1.07</b>	<b>.34</b>	<b>3.18</b>	<b>&lt;.01</b>	<b>1.40</b>	<b>.26</b>	<b>5.37</b>	<b>&lt;.001</b>
<b>Person-Level Factors<sup>c</sup></b>												
Decoding (PDE)	–	–	–	–	<b>.05</b>	<b>.02</b>	<b>3.20</b>	<b>&lt;.01</b>	<b>.05</b>	<b>.02</b>	<b>3.00</b>	<b>&lt;.01</b>
SfV	–	–	–	–	.02	.03	.62	.53	.06	.03	1.67	.09
Target Familiarity	–	–	–	–	.05	.05	.98	.32	<.01	.05	.05	.96
<b>Word-Level Factors<sup>d</sup></b>												
Log HAL Frequency	–	–	–	–	–.21	.20	–1.05	.30	–.01	.15	–.08	.94
Morphemes	–	–	–	–	–.54	.35	–1.53	.13	–.40	.26	–1.55	.12
STPT Rating	–	–	–	–	–.75	.41	–1.83	.07	–.46	.31	–1.49	.14

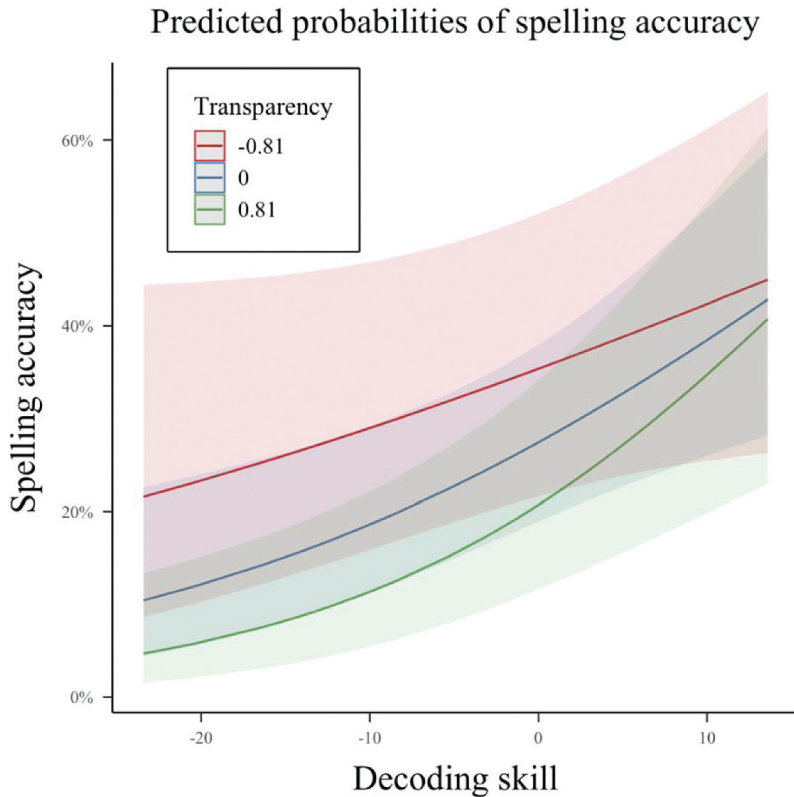
Note. Est.= Parameter Estimate; SE= Standard Error; PDE= Phonemic Decoding Efficiency; STPT= Spelling to Pronunciation Transparency; SfV= Set for Variability; CSS= Correctly Spelled Schwas. Each of the predictors and respective estimates represent the results from predicting probability of word spelling accuracy from all variables simultaneously (i.e., in the presence of all other word- and person-level predictors in the model). <sup>a</sup>Interactions are between a person-level predictor and a word-level predictor for a specific spelling item response; <sup>b</sup>Item-specific factors represent item-specific performance; <sup>c</sup>Person-level factors represent aggregate performance by the individual on the measure; <sup>d</sup>Word-level factors represent fixed characteristics of the 25 target spelling words that remain constant across individual participants.

represents an individual's ability to correctly spell one of the most difficult sublexical units of each word (i.e., schwa). These models demonstrated how even correctly spelling the most difficult-to-represent sublexical units in a word does not guarantee that one's representation for the whole word is of high enough quality to recall each letter precisely and in the correct order.

To counter this dependency issue, we ran a second exploratory interaction model with the proportion of correctly spelled schwas removed. Results from Exploratory Model II indicated a significant person-by-word interaction between decoding skill and spelling-to-pronunciation transparency rating, graphically depicted in Figure 2. Findings indicated that adults on the higher end of decoding skill have a significantly higher likelihood of correctly spelling a word, compared to peers who are poorer decoders, irrespective of transparency rating. In contrast, less skilled decoders' probability of correct spelling was notably influenced by spelling-to-pronunciation transparency, with nearly 35% lower likelihood of accurate spelling for words with lower transparency compared to highly skilled decoders and approximately 30% lower likelihood for words that were rated as being closer to the average transparency rating. These large differences in likelihood of accurate spelling suggest that, as hypothesized, poor decoders were more reliant on transparent orthographic-phonological relationships for accurate word spelling than peers who have stronger decoding skills.

## Discussion

This study's purpose was to test five main predictions made by the LQ hypothesis related to representational quality in skilled readers by modeling person-level, word-level, item-specific, and person-by-word predictors of individual differences in adult spelling performance. In support of our first prediction about variation existing across person and word in LQ of lexical representations, we observed both a wide range of spelling accuracy across words for a given individual (i.e., 8.89 words correctly spelled on average with an SD of 4.95 words) and a diverse set of errors for a given word across individuals. These results suggest that the adults in this sample varied greatly in both their representational quality of these words and knowledge sources that may have been used to fill the gaps in the orthographic representations stored for these words. To better characterize the range of spelling errors of various target words, we provided a visualization of unique errors for the most frequently



**Figure 2.** Interaction of total decoding score and spelling-to-pronunciation transparency rating in likelihood of accuracy on target spelling task. *Note.* Decoding ability and transparency rating are both mean-centered predictors with each unit away from the mean representing one standard deviation (e.g., 0.81 [red line] in the legend indicates a spelling-to-pronunciation transparency rating that is one standard deviation above the average score; -10 on the x-axis indicates a raw total score on the phonemic decoding efficiency task that is 1 standard deviation below the average raw total score).

misspelled word *assiduous*, compared to *kaleidoscope*, the fourth most difficult word to spell (each of which had 30 unique misspellings) across the dimensions of familiarity, orthographic distance, and phonological distance of the misspelling from the target (see Figure 1). For *assiduous*, which was mostly unfamiliar to the sample, participants who were familiar with the word did not stray far from the target word in terms of orthographic or phonological distance (top left panel). However, for those who were unfamiliar (top right panel), the variation in misspellings as a function of both orthographic and phonological distance from the correct spelling suggests that an individual's lower-quality representation of *assiduous* could take several forms, ranging from very impoverished (e.g., "oquious," "esidious") to somewhat recognizable (e.g., "asiduou," "esiduou"). For a word with greater overall familiarity, it is interesting to see such a wide range of orthographic distance in the misspellings of participants who were familiar with *kaleidoscope* (bottom left panel), yet a similar range of phonological distance in the misspellings compared to those who were familiar with and misspelled *assiduous* (top left panel). Results from the visualization of spelling errors associated with *assiduous* and *kaleidoscope* demonstrate that familiarity is likely necessary for correct spelling, but certainly not sufficient for building a high-quality representation. Overall, the visualization of these misspellings highlights the extreme variability across individuals in their ability to represent the spellings of low-quality representations and certainly warrants a more systematic analysis of unique spelling errors across the entire sample of target words in a future study.

Our second prediction was that measures representing an individual's level of foundational resources should make a contribution to the probability that a word is spelled correctly. This was supported in Main Effects Model I, where general English decoding and SfV were predictive of item-level spelling accuracy, along with Main Effects Model II where general decoding was still a significant predictor above and beyond item-specific predictors. These results support two important notions that follow from the LQ hypothesis. The first notion is that knowledge of orthographic-phonological relationships (as measured by decoding skill) is critical to helping one develop precise orthographic representations (Ehri, 2015; Nation, 2017; T. Ocal & Ehri, 2017a), likely supporting the tuning of items from low to high quality over time through multiple decoding opportunities and exposure to text (see Andrews, 2012; Castles et al., 2007; Hersch & Andrews, 2012). The second notion is that when presented with the challenge of spelling a word that is represented with lower quality in the lexicon, stronger decoders likely have an advantage over less skilled decoders in the ability to supplement the recall of a lower quality representation of a given word (e.g., *nauseous*) with other foundational resources. For example, a better decoder may be more flexible in considering different mappings from phonology to orthography to arrive at a plausible word spelling (e.g., "nauschious") or retrieve a second stored pronunciation of a word (i.e., redundant phonological representation) to recall the precise letters that map onto the correct pronunciation of the word (e.g., /naw-see-ous/). We acknowledge, however, that Main Effects Model I leaves a large portion of the variance in spelling accuracy to be explained by other characteristics of the individuals not measured in our study, such as print exposure (Andrews et al., 2020; Burt & Fury, 2000; Falkauskas & Kuperman, 2015) and rich encounters with words across diverse contexts (i.e., lexical legacy; Nation, 2017).

The lack of significant contributions from any of the word-level predictors in Main Effects Model I does not support our third prediction about the impact of word-level features on probability of correct spelling. Our sample of spelling words, however, was limited to a lower range of frequency (3.219 to 10.498) and did not include the lowest (1) or highest (6) possible spelling-to-pronunciation transparency rating. These restrictions in range may explain why neither frequency nor spelling-to-pronunciation transparency emerged as significant predictors of spelling accuracy. Evidence from lexical decision studies show that effects of irregular word pronunciation should be specific to lower frequency words (Seidenberg et al., 1984; Waters & Seidenberg, 1985), so our third hypothesis may only be supported with our sample of lower frequency words for individuals who struggle with forming high quality lexical representations (i.e., poorer readers) and rely on other features of a word, such as the spelling-to-pronunciation transparency, to spell correctly.

Results from Main Effects Model II support our fourth prediction about significant contributions from item-specific predictors to likelihood of spelling accuracy, showing that an individual was more likely to spell a word correctly if they a) were familiar with the target word's spoken form prior to the testing session, b) were able to correctly identify the target word from hearing its mispronounced form, or c) spelled the letters representing the schwa(s) within the target word correctly. Of key interest, a significant contribution from item-specific SfV supports the LQ assumption that formation of a high-quality representation also relies on storing redundant phonological representations, including at least one that is recoverable from regular orthographic-to-phonological mappings (see A. Edwards et al., 2021; Elbro, 1998; Elbro & Jensen, 2005; Goswami, 2000; C. Perfetti & Hart, 2001). Given that the mispronunciations participants heard during the SfV task represented decoded forms of the words in the target spelling task, we interpret this result to mean that when more highly skilled spellers hear a word's correct pronunciation in the spelling task, they may activate a second plausible pronunciation of that word based on orthographic-to-phonological mappings (i.e., a redundant pronunciation similar to the decoded form presented in the SfV task) to aid in recalling the specific graphemes that can represent more ambiguous phonemes in the word's correct pronunciation. We argue that storage of this second pronunciation may be a byproduct of the complete decoding that is crucial for successful self-teaching of new words (Share, 1995, 2008) and formation of high-quality orthographic representations (Nation & Castles, 2017; C. A. Perfetti, 1992). Successfully recognizing a word from its mispronunciation in the SfV task, therefore, can be understood as

a measure of redundancy for a specific lexical representation, which is necessary, but not sufficient, for accurate spelling.

Alternatively, one may consider that the ability to accurately recall a word's spelling can increase the likelihood of identifying a word from its mispronunciation in the SfV task by actively mapping a mispronunciation onto a letter string, recognizing it as being similar or equivalent to the spelling of a known word, and correctly pronouncing the word in response. However, given the length, generally low frequency, and irregularity of the target words, it is unlikely that individuals were actively engaging in this conversion (mispronunciation to letter string to correct pronunciation) for each familiar target word, especially since the task included a prompt to respond within 10 seconds of hearing the mispronunciation before hearing the next item's mispronunciation. We speculate that characteristics of these irregular words and the task's administration, paired with the participants' lack of familiarity with the task, likely created a high cognitive load that would make it very difficult for a person to engage in this effortful recoding strategy and provide a correct response within the allotted response time. Future investigations may determine which of these two explanations is supported in adult participants by explicitly asking participants to report item-specific strategy use throughout the SfV task (e.g., immediate visualization of a word that matches the decoded form heard in the task versus actively recoding the decoded form to a plausible spelling and back to a known word's pronunciation during the task).

Item-specific contributions from SfV to spelling accuracy may also be explained as skilled spellers simply performing better in SfV because they are generally more flexible with pronunciations of English words without the need to store and actively recall alternate pronunciations for specific words. Nevertheless, we believe this explanation is also unlikely given that general SfV performance, which included performance on higher frequency non-target spelling items in the total accuracy score (i.e., 40 items instead of 25), only emerged as a significant predictor in Main Effects Model I, but none of the subsequent models when item-specific SfV performance was modeled simultaneously. This deviates from recent reports of SfV, both general (i.e., person-level) and item-specific, being simultaneously predictive of word reading performance in younger developing readers (Steady, Compton, et al., 2019). Importantly, this difference in the roles of general and item-specific SfV in spelling versus reading performance supports the LQ assumption that accurate spelling relies heavily on high-quality lexical entries that are only influenced by highly constrained sources within the lexicon, such as a stored mispronunciation for a given word, compared to lower-quality lexical entries that permit multiple sources of information outside of lexical constraints, such as a metalinguistic strength in phonological flexibility that is impacted by overall (rather than item-specific) reading experience.

Next, the significant contribution from proportion of correctly spelled schwas to likelihood of accurate spelling highlights the importance of one's ability to recall precise phonological to orthographic connections that are ambiguous, like the schwa, and unrecoverable from simply using frequent orthographic-phonological relations. This aligns with Block and Duke's (2015) report of schwa vowels being the most difficult sound to spell for young children and T. Ocal and L. C. Ehri's (2017b) findings from a training study focused on improving adults' spelling of difficult English words containing letters that did not map directly onto sounds, including schwas. Our findings here support the LQ assumption that even skilled readers can have plenty of low-quality representations, and therefore, struggle with accurately spelling words that are familiar from previous encounters in reading (C. A. Perfetti, 1992) and contain hard-to-map phonemes, like the schwa. Relatedly, a significant contribution from item-specific familiarity to likelihood of spelling accuracy supports C. Perfetti's (2007) notion that formation of a high-quality representation relies on multiple encounters with a given word to strengthen the connections between orthographic, phonological, and semantic knowledge of that word.

Finally, support for our fifth prediction is interpreted from the last set of exploratory interaction models. Given that words containing schwas are more likely to be rated as less transparent in spelling to pronunciation (A. Edwards et al., 2021) and the importance of strong decoding skill in spelling



performance (T. Ocal & Ehri, 2017a), we speculated that in Exploratory Model I, the interaction of decoding skill by word transparency rating may be competing with proportion of correctly spelled schwas for very similar amounts of variance in likelihood of spelling accuracy. Thus, when we removed proportion of correctly spelled schwas from Exploratory Model II, it was interesting that the interaction significantly predicted likelihood of spelling accuracy, supporting our hypothesis that poorer decoders would be more reliant on more transparent and frequent orthographic-phonological relationships to spell a word correctly, including spelling of the schwa. This interaction also offers more nuanced support for our third prediction about word-level characteristics contributing to the probability of correct spelling by highlighting the strong impact of lower spelling-to-pronunciation transparency on poor decoders' spelling accuracy, whereas highly skilled decoders' probability of correct spelling is seemingly unaffected by a word's spelling-to-pronunciation transparency rating (see Waters et al., 1984).

Considering the 30% base rate of spelling accuracy for the entire sample, there are at least two possibilities for why highly skilled decoders' probability of correct spelling is significantly less impacted by spelling-to-pronunciation transparency than that of poorer decoders. One possibility is that these strong decoders tend to have more high-quality representations that are activated for accurate spelling as a result of more successful decoding attempts in prior reading encounters with the target words they were able to recall and spell correctly. This interpretation would support the notion that skilled readers' continuous "tuning" (see Andrews, 2012; Castles et al., 2007; Hersch & Andrews, 2012) of lexical representations through reading experience is responsible for the additions of high-quality representations to the lexicon over time that allow for autonomous recognition of a word from its spoken form and accurate recall of its precise written form. This would also expand C. A. Perfetti and Hart's (2002) conclusion about skilled readers likely drawing from a more coherent lexical knowledge structure in reading tasks (compared to less skilled readers) to the task of spelling. In other words, understanding the close links between orthographic and phonological structures at multiple levels within a word is critical for both reading *and* spelling.

Another interpretation takes into account the wide range of familiarity with the target spelling words, including 2 participants who were only familiar with 10 out of 25 target spelling words. Following the LQ hypothesis, if skilled readers' lexical representations vary in quality across individual and word, one may conclude that while grappling with the activation of a low-quality representation after hearing a word in the spelling task, strong decoders are better equipped to supplement their lower-quality representations for an unfamiliar word with other resources (i.e., stored mispronunciation or active sound-to-letter conversion) compared to poor decoders who are likely also depending on lower-quality entries for unfamiliar words, but without the additional resources (i.e., knowledge of decoding rules, flexibility with less frequent orthographic-phonological relationships) to supplement their representations.

### **Limitations**

One limitation that must be acknowledged is our study's lack of both an item-specific and general vocabulary measure. In our attempts to conceptualize total familiarity as a proxy for phono-semantic connections, we recognize that this measure is not a sufficient replacement for general English vocabulary or item-specific semantic knowledge (Balota & Spieler, 1999). Moreover, adapting a receptive vocabulary measure, such as the PPVT-IV (Dunn & Dunn, 2007), or an expressive vocabulary measure, such as the Woodcock-Johnson III Picture Vocabulary (Woodcock et al., 2001) to test participants' knowledge of item-specific meanings would be critical in testing LQ assumptions regarding the importance of combining information about a word's meaning and contexts of use with the phonological and orthographic knowledge that are stored for that word. Additionally, it must be noted that due to the design of the spelling task (i.e., participant heard the word, attempted the spelling, and then indicated if they were (un)familiar with the word prior to the testing session), the participants' responses about their item-specific familiarity for a given word may have been influenced

by their confidence in the accuracy of their attempted spelling for that word. In future studies, investigators should consider administering a separate familiarity measure with non-target foil words mixed in the list of target words, either before participants encounter the words in the dependent measure of interest or in a separate testing session.

## Conclusion

This study used item-level spelling performance in adult skilled readers to test a set of hypotheses derived from Perfetti's LQ hypothesis as it relates to the formation of high-quality lexical representations. Overall, results supported the hypotheses and illustrated the utility of using spelling as an indicator of LQ within individuals. We found significant variation with respect to representational quality, supporting C. A. Perfetti and Hart's (2002) assertion that LQ varies across individuals for a specific word and across words for a specific individual. Item-specific familiarity, proportion of correctly spelled schwas, and SfV were associated with the presence of high-quality representations within individuals. Item-specific SfV was a particularly interesting predictor, suggesting that a second plausible pronunciation of the word derived from applying orthographic-to-phonological mappings is important to activate in the formation of high-quality lexical representations. This is consistent with item-level orthographic learning theories stipulating that fully decoding a word is necessary for the formation of high-quality representations (Nation & Castles, 2017; C. A. Perfetti, 1992; Share, 1995). We also reported an interaction that expands the LQ hypothesis predictions for reading to spelling performance, where an individual's production of high-quality lexical representations depends on both general decoding ability (person-level contribution) and the distance between a word's correct and decoded pronunciations (word-level contribution). We believe our results in adult participants should encourage more empirical investigations of item-level spelling performance in samples of children to help illuminate processes contributing to the formation of high-quality representations in developing readers.

Furthermore, our findings support T. Ocal and L. C. Ehri's (2017b) success in training adult poor spellers to use an overpronunciation (i.e., spelling pronunciations; redundant phonological representation) strategy for improved spelling accuracy between pre- and posttest. Significant improvements from just a brief training are promising in terms of better understanding how to support students who are struggling to form high-quality lexical representations with strategies that directly and quickly strengthen orthography-to-phonology connections. Given our finding of the decoding by spelling-to-pronunciation transparency interaction in spelling accuracy, future efforts in testing the effects of an overpronunciation training strategy with attention to flexible orthography-to-phonology mappings for opaque sublexical units should also consider the importance of the variation in words' spelling-to-pronunciation transparency reported in the current study. These efforts may lead to more concrete recommendations for spelling instruction that simultaneously supports reading development of individuals across the range of general decoding and spelling ability, including struggling readers.

## Notes

1. In the spirit of developing ideas that work across different complementary theoretical frameworks, and in service of a more descriptive account of the phenomena dealt with in this paper, we intentionally use several terms in ways that are theory agnostic. The terms "lexical representation" and "word representation" are used interchangeably, where the term "lexicon" is intended only to convey that knowledge about words is represented in the mind, rather than making a specific commitment about the manner in which that representation takes place. In terms of the nature of the representations, we hold only that representational information includes, at least, featural information about orthography, phonology, and semantics, and assert that this view is consistent both with theories about lexical quality and other theories about representation learning related to reading (e.g., Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989). To this end, the term "storage" is only intended to convey something general about the presence of featural information in the mind about these knowledge domains, and

the activation of this information in service of reading and spelling processes. Specifications about the nature of the representational system (i.e., how representations are encoded in the cognitive system) are important but outside the scope of this work.

2. We consider an item to be a high-quality representation within a person when the word is correctly spelled; whereas words spelled incorrectly are considered lower-quality items.
3. One manual correction to “medieval” was completed by an experienced speech language pathologist with consideration of the mainstream English dialect of General American English.

## Acknowledgments

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 Development (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. The authors thank James Elliot (lab manager) and the research assistants who were instrumental in data collection/entry: Daniel Abes, Logan Bell, Katherine Diaz, Cristina Himelhoch, Kara Muston, Yohana Pino, and Rebecca Vasile. Additional thanks to Dr. Sonia Cabell and Dr. Keisey Fumero for their input on this work. No potential conflict of interest was reported by the author(s).

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The work was supported by the Institute of Education Sciences [R324B190025]; National Institute of Child Health and Human Development [P20HD091013].

## ORCID

Valeria M. Rigobon  <http://orcid.org/0000-0002-7089-1764>  
 Nuria Gutiérrez  <http://orcid.org/0000-0002-2809-1608>  
 Ashley A. Edwards  <http://orcid.org/0000-0002-5268-4235>  
 Nancy Marencin  <http://orcid.org/0000-0002-5457-972X>  
 Matt Cooper Borkenhagen  <http://orcid.org/0000-0002-8245-0130>  
 Laura M. Steacy  <http://orcid.org/0000-0003-1603-2990>  
 Donald L. Compton  <http://orcid.org/0000-0003-3855-3746>

## References

- Andrews, S. (2012). Individual differences in skilled visual word recognition and reading: The role of lexical quality. In J. S. Adelman (Ed.), *Visual word recognition* (Vol. 2, pp. 151–172). Psychology Press.
- Andrews, S., & Treiman, R. (2015). Individual differences among skilled readers: The role of lexical quality. In A. Pollatsek (Ed.), *The oxford handbook of reading; the oxford handbook of reading* (pp. 151–174). Oxford University Press.
- Andrews, S., Veldre, A., & Clarke, I. E. (2020). Measuring lexical quality: The role of spelling ability. *Behavior Research Methods*, 52(6), 2257–2282. <https://doi.org/10.3758/s13428-020-01387-3>
- Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, 128(1), 32. <https://doi.org/10.1037/0096-3445.128.1.32>
- 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(3), 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>
- Block, M. K., & Duke, N. K. (2015). Letter names can cause confusion and other things to know about letter-sound relationships. *Young Children*, 70(1), 84–91.

- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, 45(2), 122–159. <https://doi.org/10.1080/01638530701792826>
- Bosman, A. M., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. E. Fayol (Eds.), *Learning to Spell: Research, Theory, and Practice Across Languages* (Vol. 10, pp. 173–194). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological methods*, 23(3), 389.
- Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, 40(3), 226–243. <https://doi.org/10.1177/00222194070400030401>
- Burt, J. S., & Fury, M. B. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing*, 13(1/2), 1–30. <https://doi.org/10.1023/A:1008071802996>
- Caruana, N., Colenbrander, D., & McArthur, G. (2019). *The Macquarie University Advanced Adult Spelling Test (MAAST)*. Macquarie Online Test Interface.
- Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, 97(3), 165–182. <https://doi.org/10.1016/j.jecp.2007.01.006>
- 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>
- Cho, S. J., Partchev, I., & De Boeck, P. (2012). Parameter estimation of multiple item response profile model. *British Journal of Mathematical and Statistical Psychology*, 65(3), 438–466. <https://doi.org/10.1111/j.2044-8317.2011.02036.x>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th Ed.). Pearson.
- Edwards, A., Rigobon, V. M., Steacy, L., & Compton, D. (2021, November 28). Spelling to pronunciation transparency ratings. <https://doi.org/10.31234/osf.io/2wmk5>
- Edwards, A. A., Steacy, L. M., Siegelman, N., Rigobon, V. M., Kearns, D. M., Rueckl, J. G., & Compton, D. L. (2021). Unpacking the unique relationship between set for variability and word reading development: Examining word- and child-level predictors of performance. *Journal of Educational Psychology*, 114(6), 1242–1256. <https://psycnet.apa.org/doi/10.1037/edu0000696>
- Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, mostly. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237–270). Erlbaum.
- Ehri, L. C. (2015). How Children Learn to Read Words. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading* (pp. 293–325). New York, NY: Oxford University Press.
- Elbro, C. (1998). When reading is “readn” or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology*, 39(3), 149–153. <https://doi.org/10.1111/1467-9450.393070>
- Elbro, C., & de Jong, P. (2017). Orthographic learning is verbal learning: The role of spelling mispronunciations. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 148–168). John Benjamins. <https://doi.org/10.1075/swll.15.10elb>
- Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A. (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., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, 46(4), 375–384. <https://doi.org/10.1111/j.1467-9450.2005.00468.x>
- Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 41(6), 1607–1627. <https://doi.org/10.1037/xlm0000132>
- Gibson, E. J., & Levin, H. (1975). *The psychology of reading*. The MIT press.
- Goswami, U. (2000). Phonological and lexical processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 251–267). Lawrence Erlbaum Associates Publishers.
- 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. <https://doi.org/10.1037/0033-295X.111.3.662>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Herscht, J., & Andrews, S. (2012). Lexical quality and reading skill: Bottom-up and top-down contributions to sentence processing. *Scientific Studies of Reading*, 16(3), 240–262. <https://doi.org/10.1080/10888438.2011.564244>
- Holmes, V. M., & Malone, N. (2004). Adult spelling strategies. *Reading and Writing*, 17(6), 537–566. <https://doi.org/10.1023/B:READ.0000044368.17444.7d>
- Hsiao, Y., & Nation, K. (2018). Semantic diversity, frequency, and the development of lexical quality in children’s word reading. *Journal of Memory and Language*, 103, 114–126. <https://doi.org/10.1016/j.jml.2018.08.005>

- 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>
- Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *Science of Learning*, 2(1), 1–4. <https://doi.org/10.1038/s41539-017-0004-7>
- 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>
- Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*, 69(4), 996–1011.
- Ocal, T., & Ehri, L. (2017a). Spelling ability in college students predicted by decoding, print exposure, and vocabulary. *Journal of College Reading & Learning*, 47(1), 58–74. <https://doi.org/10.1080/10790195.2016.1219242>
- Ocal, T., & Ehri, L. C. (2017b). Spelling pronunciations help college students remember how to spell difficult words. *Reading and Writing*, 30(5), 947–967. <https://doi.org/10.1007/s11145-016-9707-z>
- Ormrod, J. E., & Jenkins, L. (1989). Study strategies for learning spelling: Correlations with achievement and developmental changes. *Perceptual and Motor Skills*, 68(2), 643–650. <https://doi.org/10.2466/pms.1989.68.2.643>
- 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. (1985). *Reading ability*. Oxford University Press.
- Perfetti, C. A. (1991). The psychology, pedagogy, and politics of reading. *Psychological Science*, 2(2), 70–83.
- Perfetti, C. A. (1992). The representation problems in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Erlbaum. <https://doi.org/10.4324/9781351236904-6>
- 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). John Benjamins Publishing Company. <https://doi.org/10.1075/z.206.04per>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). John Benjamins Publishing Company. <https://doi.org/10.1075/swll.11.14per>
- Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection* (pp. 67–86). American Psychological Association. <https://doi.org/10.1037/10459-004>
- Perfetti, C., & 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>
- R Development Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. ISBN 3-900051-07-0, ISBN 3-900051-07-0. <http://www.R-project.org/>
- Rigobon, V. M., Gutiérrez, N., Edwards, A. A., Abes, D., Steacy, L. M., & Compton, D. L. (2023). Does Spanish knowledge contribute to accurate English word spelling in adult bilinguals? *Bilingualism: Language and Cognition*, 1–18. <https://doi.org/10.1017/S1366728923000093>
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523. <https://doi.org/10.1037/0033-295X.96.4.523>
- Seidenberg, M. S., Waters, G. S., Barnes, M., & 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/S0022-53718490270-6>
- 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](https://doi.org/10.1016/0010-0277(94)00645-2)
- Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 36, pp. 31–82). Elsevier.
- Stafura, J., & Perfetti, C. (2017). Integrating word processing with text comprehension: Theoretical frameworks and empirical examples. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 9–32). John Benjamins Publishing Company. <https://doi.org/10.1075/swll.15.02sta>
- Stacy, L. M. (2020). Capitalizing on the promise of item-level analyses to inform new understandings of word reading development. *Annals of Dyslexia*, 70(2), 153–159. <https://doi.org/10.1007/s11881-020-00203-z>
- Stacy, L. M., Compton, D. L., Petscher, Y., Elliott, J. D., Smith, K., Rueckl, J., Sawi, O., Frost, S., & Pugh, K. (2019). Development and prediction of context-dependent vowel pronunciation in elementary readers. *Scientific Studies of Reading*, 23, 49–63. <https://doi.org/10.1080/10888438.2018.1466303>
- Stacy, L. M., Edwards, A., Rigobon, V. M., Gutierrez, N., Marencin, N. C., Siegelman, N., Himelhock, A., Himelhock, C., Rueckl, J., & Compton, D. L. (2022). Set for variability as a critical predictor of word reading: Potential implications for early identification and treatment of dyslexia. *Reading Research Quarterly*, 58(2), 254–267. <https://doi.org/10.1002/rrq.475>

- Steady, L. M., Kearns, D. N., Gilbert, J. K., Compton, D. L., Cho, E., Lindstrom, E. R., & Collins, A. A. (2017). Exploring individual differences in irregular word recognition among children with early-emerging and late-emerging word reading difficulty. *Journal of Educational Psychology, 109*, 51–69. <https://doi.org/10.1037/edu0000113>
- Steady, L. M., Wade Woolley, L., Rueckl, J. G., Pugh, K. R., 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>
- Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of word reading efficiency 2*. Pro-Ed.
- 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>
- Van Dyke, J. A., & Shankweiler, D. P. (2012). From verbal efficiency theory to lexical quality. In A. Britt & S. Goldman, & J.-F. Rouet (Eds.), *Reading-from words to multiple texts* (pp. 115–132). Routledge.
- Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition, 13*(6), 557–572. <https://doi.org/10.3758/BF03198326>
- 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*(3), 293–305. <https://doi.org/10.3758/BF03197678>
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III NU complete*. Riverside Publishing.