



# HHS Public Access

Author manuscript

*J Exp Child Psychol.* Author manuscript; available in PMC 2019 March 25.

Published in final edited form as:

*J Exp Child Psychol.* 2019 February ; 178: 226–250. doi:10.1016/j.jecp.2018.09.007.

## Examining the Role of Imageability and Regularity in Word Reading Accuracy and Learning Efficiency Among First and Second Graders At-Risk for Reading Disabilities

Laura M. Steacy and

Florida Center for Reading Research, Florida State University, (lsteacy@fcrr.org)

Donald L. Compton

Florida Center for Reading Research, Florida State University

### Abstract

This paper explores word reading accuracy and word learning efficiency in first and second grade students ( $N=125$ ). In two experiments, students participated in a short training exposing them to words that varied on orthographic regularity and imageability. In experiment 1 the form of word feedback was manipulated (phonemic vs. whole word) whereas in experiment 2 pretraining exposure was manipulated (training on meaning vs. imageability). Crossed-random effects models were used to explore child- and item-level predictors related to number of exposures to mastery, posttest word reading performance, and maintenance performance after 1-week. Results from experiment 1 indicate that imageability plays a role in irregular word learning. Results from experiment 2 suggest that students who received imageability training required fewer exposures to reach mastery. There was a significant interaction between initial word reading skill and condition, with students with low word reading skills requiring fewer exposures for mastery if they were in the imageability condition. Overall, these findings suggest that word imageability significantly impacts both word reading accuracy and rate of word learning.

### Keywords

Word reading; regularity; imageability

---

Accurate and automatic word reading provides the foundation for success in reading (Adams, 1994; Foorman & Torgesen, 2001; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001) with individual differences in word recognition skill contributing to reading comprehension variance across the entire school-age range (e.g., Johnston, Barnes, & Desrochers, 2008; García, & Cain, 2014; Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011). Typically developing students learn to read words with relatively few exposures (Reitsma, 1983), whereas students with reading disabilities typically require many more exposures to learn a word while also retaining less complete representations of words (Ehri, 1997; Ehri & Saltmarsh, 1995). According to Ehri (1995; 2014), skilled readers successfully map all graphemes in a word onto all phonemes in pronunciations creating a consolidated representation, whereas poor readers tend to form partial connections between

---

graphemes and phonemes that are less consolidated and do not lead to automatic access to the phonological and semantic representations. This leads poor readers to rely on other sources of information (e.g., local text content), that are considerably less efficient, to facilitate word recognition (see Stanovich, 1980).

Much of the early work on word recognition development has focused on the link between orthography and phonology, with less attention paid to the role of semantics. Recently, however, the role of semantics in facilitating the developmental of word recognition skills has begun to receive greater attention. Much of this work is grounded within the lexical quality hypothesis (Perfetti, 2007), which proposes that lexical representations of words, both within and across individuals, vary in the extent and strength with which aspects of their form (phonology, morphosyntax, orthography) and meaning (semantics) are represented. Rapid and automatic word identification in skilled readers has been shown to be dependent on high-quality lexical representations (see, Perfetti & Stafura, 2014; Taylor & Perfetti, 2016).

Computational models of word reading development have, to varying degrees, incorporated aspects of the lexical quality hypothesis into their architecture. For instance, the triangle (e.g., Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989), dual route (e.g., Coltheart, 2006) and connectionist dual process (Perry, Ziegler, & Zorzi, 2007; Pritchard, Coltheart, Marinus, & Castles, 2018; Ziegler, Perry, & Zorzi, 2014) approaches to visual word recognition allow two “routes” from orthography to semantics – a direct route and an indirect route via phonological recoding. While aspects of the models differ, all three models allow for a “division of labor” such that irregular words rely more heavily on the operation of mapping from orthography to semantics (see Grainger & Ziegler 2011; Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996). Orthographies such as English present unique challenges to readers, when compared to more orthographically consistent languages like Greek and Finnish, due to its reduced consistency (Seymour, Aro, & Erskine, 2003).<sup>1</sup> Thus, phonological recoding of unfamiliar words in English often results in a mismatch between pronunciation and the phonological representation stored in the lexicon, requiring semantic feedback to clean up the mismatch (see Dyson, Best, Solity, & Hulme, 2017). A limited but expanding behavioral literature supports this concept of a division of labor in which semantics support the learning of irregular words in opaque orthographies such as English (Elbro, de Jong, Houter, & Nielsen, 2012; Keenan, & Betjemann, 2007; Nation, & Cocksey, 2009; Ricketts, Nation, & Bishop, 2007). The purpose of this study was to further explore how child and item features related to phonology, orthography, and semantics impact word reading and rate of word learning in developing readers within a controlled experimental design.

## The Role of Exposure and Semantics in Word Reading Development

Increasing evidence suggests that readers rely at least partially on word meaning to aid in the development of word-specific links between orthography and phonology (see Ouellette,

---

<sup>1</sup>Consistency here refers to the degree to which phonemes map directly onto graphemes in words. In English, the mapping of orthography to phonology falls along a continuum with some irregular words mappings more directly onto phonology than others (e.g. *pint* and *touch* map more directly onto their phonological forms than *yacht* and *suede*).

2006). According to Nation and Cocksey (2009), familiarity with a word's phonological form (also known as lexical phonology) is a significant predictor of word reading, particularly for irregular words. Additionally, Taylor, Plunkett, and Nation (2011) found that pre-exposure to item definitions increased word learning of an artificial orthography in adults. Ouellette and Fraser (2009) also found that presenting words along with semantic information resulted in significantly more accurate word recognition than presenting words in isolation. Support for the importance of semantics also comes from computational models of reading (e.g. Plaut, 1998; Plaut and Shallice, 1993). Specifically, connectionist models suggest that the addition of a semantic processor (represented as item-specific knowledge) to a model containing phonological and orthographic processors improves both nonword and irregular word recognition (see Plaut, McClelland, Seidenberg, & Patterson, 1996). Furthermore, Ricketts, Nation, and Bishop (2007) found that item-specific vocabulary knowledge accounted for unique variance in irregular word reading in developing readers. Having item-specific vocabulary knowledge for a word has also been shown to be a significant predictor of orthographic learning within a self-teaching model of reading development (Wang, Nickels, Nation, & Castles, 2013). Keenan and Betjemann (2007) have speculated that item-specific semantic activation may help to "fill voids" in phonological-orthographic processing in individuals with poor mappings, such as children with reading disabilities (p. 193).

More recent studies have found that semantic knowledge of words influences both regular and irregular words in English (see Ricketts, Davies, Masterson, Stuart, & Duff, 2016) and word reading in Spanish (Davies, Barbon, & Cuetos, 2013), a transparent language with no irregularities. These findings differ from the findings of others, who found more of an impact of semantics on irregular word reading (Nation & Cocksey, 2009; Ricketts, Nation, & Bishop, 2007). The authors attribute this difference in results to the use of more advanced modeling techniques in the form of linear mixed-effects models (another word for the crossed random effects models used in the current study) that allow word-specific semantic knowledge to predict word-specific reading ability (e.g., does having knowledge of what the word *yacht* means influence a child's ability to read the word *yacht*). The current study adds to this evolving literature by employing item-based item response theory (IRT) models (i.e., crossed random effects) to explore the role of imageability, a semantic feature, on word reading accuracy and learning efficiency of regular and irregular words in developing readers.

## The Role of Regularity and Imageability in Word Reading

### Word regularity.

Word regularity refers to the degree to which the pronunciations of phonemes within a word reflect common spelling-sound correspondences (Metsala, Brown, & Stanovich, 1998). There has been consistent support for the important role regularity plays for both typically developing students and students with or at-risk for reading disabilities (Balota & Ferraro, 1993; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Stanovich & Bauer, 1978; Waters & Seidenberg, 1985). Despite the importance of regularity in word reading development, there remains considerable variance to explain at the word level, suggesting that there may be

other important word characteristics to consider (Griffiths & Snowling, 2002; Steady et al., 2017; Wang, Nickels, Nation, & Castles, 2013).

### **Imageability.**

Imageability is a word specific semantic feature that refers to the ease with which a word can elicit a mental image in the reader (Paivio, Yuille, & Madigan, 1968). An example of a high imageability word is *farm*, while a low imageability word is *which*. This word feature has received some attention given that it predicts unique variance in word reading accuracy and lexical decision times (see Pexman, 2012). Imageability is highly correlated with word concreteness/abstractness and the words are sometimes used interchangeably in the literature. We posit, along with others (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Laing & Hulme, 1999; Pexman, 2012; Strain, Patterson, & Seidenberg, 1995), that imageability is a word feature associated with the semantic representation of the word. Researchers have postulated that imageability and concreteness contribute to how easily words are remembered because these words activate both perceptual and verbal memory codes (see Brysbaert, Warriner, & Kuperman, 2014; Paivio, 1991; Sabsevitz, Medler, Seidenberg, & Binder, 2005). Furthermore, studies suggest that imageability/concreteness can support word learning in bilingual individuals who may be more attuned to semantic information when learning to read words (see Kaushanskaya & Rehtzigel, 2012).

Evidence from studies on context-free single word reading demonstrate that semantic word features such as concreteness, imageability, and meaningfulness all play a role in both word recognition and lexical decision tasks (see e.g., Duff & Hulme 2012; Keenan & Betjemann, 2007; Laing & Hulme, 1999). For instance, previous studies in adults indicated that as irregular words become more imageable, recognition accuracies increase (Strain, Patterson, & Seidenberg, 1995) and reaction times decrease (Strain & Herdman, 1999). In children, Duff and Hulme (2012) found that when children learned words that varied in imageability and spelling sound consistency (i.e., regularity), imageability impacted word reading accuracy, particularly in later trials and for words that were irregular. Also, having item-specific vocabulary knowledge for words has been shown to be a significant predictor of orthographic learning (Wang, Nickels, Nation, & Castles, 2013) and having passage specific prior knowledge of the content of a text decreases the number of word reading errors in poor readers (Priebe, Keenan, & Miller, 2012). Furthermore, imageability has been reported to be a particularly important word feature for poor readers (Coltheart, Laxon, & Keating, 1988). Some studies examining the role of imageability in children have focused on learning nonwords. Laing and Hulme (1999) approached this issue using word abbreviations that children were encouraged to pair with spoken English words. Abbreviations for high imageability words (e.g., ltr for ladder) were learned more easily than abbreviations for low imageability words (e.g., lzn for listen). Despite these findings, semantics, and in particular imageability, has received little attention within the literature on at-risk readers and thus the role of word-level semantics requires more attention (Keenan & Betjemann, 2007).

### **Causal inferences of the role of semantics in word reading.**

To date, there has been very little work that allows us to make causal inferences about the role of semantic word features such as imageability in word reading acquisition. More

specifically, there have been no previous studies that have attempted to directly manipulate imageability to improve instruction and to establish a causal relationship between imageability and word reading. The only effort known to the authors in this area was a study on nonword learning by Duff and Hulme (2012). In this study the authors attempted to make nonwords imageable using sentences (experiment 2). They compared two conditions, one with only phonological information for nonwords and one with both phonological and semantic information. They found no added benefit for semantic knowledge (i.e., imageability) over and above the benefit of phonological information. Wang, Nickels, Nation, and Castles (2013) used a vocabulary training condition that included picture supports for general word learning. They found that vocabulary knowledge of the words, supported by pictures, was only advantageous for irregular words. This study did not, however, focus only on low imageability or abstract words. We argue that irregular real words may present unique challenges to students, particularly those with or at-risk for reading disabilities. Imageability may be one factor that facilitates the reading and learning of these words. This study focuses on these uniquely challenging words.

## The Current Study

The current study addresses important gaps within the developmental word reading literature by presenting results from two experiments exploring word reading accuracy and word reading efficiency (i.e., trials to mastery) in first and second grade children at risk for developing reading disabilities. The two experiments in this study build on each other by first focusing on the role of imageability and regularity as predictors of word reading accuracy and word learning efficiency (Experiment 1) and then attempting to improve word reading accuracy and word learning efficiency of irregular words by training word imageability (Experiment 2). In Experiment 1 we also included two feedback conditions for whole word feedback and phonological analysis feedback per the work of Ehri (1995), Spaii, Ellerman, & Reitsma (2004), van Daal and Reitsma (1990). This study expands the existing literature by: (1) exploring the role of imageability in both word reading accuracy (both at posttest and maintenance) and efficiency of word learning (i.e., the number of exposures required for mastery), (2) examining the imageability and regularity phenomenon in students with or at-risk for reading disabilities, and (3) comparing improvements in word reading accuracy and word learning efficiency across groups of children exposed to imageability training, a congruent vocabulary training, and an active control group. The design used in Experiment 2 allowed us to explore whether there is something uniquely facilitative about imageability training that is not accomplished through traditional vocabulary training.

Overall, we hypothesized that imageability would facilitate word reading, especially for irregular words. Based on previous literature, we expected students to have a lower probability of reading low imageability irregular words correctly than high imageability and regular words. We also hypothesized that students would require fewer exposures for mastery of highly imageable and regular words. Finally, in terms of training, we hypothesized that semantic support in the form of imageability and/or vocabulary training would help students to learn irregular words more efficiently.

## Experiment 1

### Method

**Participants.**—For Experiment 1, 47 at-risk children were drawn from one rural school district in the Southeastern region of the United States. Fifteen first and second grade teachers nominated five to eight students in their classes with the lowest levels of reading skill and were asked to send consent forms home with these students. We pretested all students using measures of word reading, pseudoword reading, phonological awareness, rapid naming, and vocabulary. Students were also pretested on the study target words.

#### **Child-level measures.**

**Picture vocabulary.:** The picture vocabulary test of the Woodcock Johnson-III (Woodcock, McGrew, & Mater, 2001), a measure of expressive vocabulary, required students to identify pictured objects. The test was discontinued after the student got the six highest-numbered items on a page incorrect. The manual reports a split-half reliability of .70 for six year olds and .71 for seven year olds.

**Word identification.:** The Word Identification task from the Woodcock Reading Mastery Test (WRMT-III; Woodcock, 2011) required children to read isolated real words aloud. The test was discontinued when the student incorrectly identified four consecutive items. Woodcock (2011) reported a split-half reliability of .96 and .94 for first and second grade, respectively.

**Word attack.:** The Word Attack task from the WRMT-III (Woodcock, 2011) required children to read isolated pseudowords aloud (e.g., ree, ip, and weaf). The test was discontinued after four consecutive errors. Reported split-half reliabilities were .96 and .92 for first and second grade, respectively (Woodcock, 2011).

**Phonemic awareness (PA).:** The phonemic awareness task we used in this experiment was an elision task that required students to delete phonological units from words. The task began with syllable deletion and transitioned to deletion of initial, final, and medial phonemes. The task was based on a task from the Rosner Test of Auditory Awareness Skills (TAAS; Rosner, 1979).

**Rapid automatized naming (RAN).:** We used a rapid naming task developed by Denckla and Rudel (1976). Students were given one practice item with five letters. We then presented them with a page of those letters (p, o, a, s, d) in random order and asked them to name the letters as quickly as possible.

**Target word reading.:** The target word reading task was researcher developed and required students to read a list of 32 words. These words were selected using a factorial design based on regularity and imageability. Eight words were selected to be highly imageable regular words, eight to be highly imageable irregular words, eight to be regular low imageability words, and eight to be irregular low imageability words. These words are provided in Table 1. We controlled for frequency, number of letters, and initial phoneme when selecting the

words such that no significant differences on these dimensions existed between word classes. Analyses of variance indicate that there were no differences between the four classes of words based on orthographic neighborhood size,  $F(3, 28)=.57, p=.64$ , phonological neighborhood size,  $F(3, 28)=.56, p=.65$ , or bigram frequency by position,  $F(3, 28)=1.90, p=.15$ . The target word reading measure was administered at pretest, posttest (i.e., post-training), and maintenance. Descriptive statistics for the target words are provided in Appendix A.

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

**Imageability.:** Imageability is a word specific feature referring to the ease with which a word can elicit a mental image in the reader (Paivio, Yuille, & Madigan, 1968). We used the data available from the MRC Psycholinguistic databases (imageability scores based on Paivio, Yuille and Madigan (1968), Toglia and Battig (1978) and Gilhooly and Logie (1980)). Three items did not have data available from this source. We used updated imageability ratings by Bird, Franklin, and Howard (2001) for those items. The list of words they used for their study ( $N= 2645$ ) contained 75 items that also had imageability ratings in the MRC database. They report a correlation of .92 ( $p<.001$ ) between their ratings and the MRC ratings.

**Regularity.:** Words were considered irregular if they were not consistent with typical letter sound correspondences. We based the regularity criteria on the coding system offered by Rastle and Coltheart (1999).

**Experimental design.:** The design of this experiment was based on a design used in word learning studies conducted by Martin-Chang and Levy (2005; 2006) and Martin-Chang, Levy, and O'Neill (2007). This experiment was conducted over the course of a two-week period. Students first participated in a brief pretesting session (30 minutes) on the first day, three training sessions, with four exposures per word each day on three consecutive school days (15-20 minutes), one posttesting session on the fifth day (15-20 minutes), and maintenance testing one week after posttesting. Students were rank ordered based on pretest performance and randomly assigned to whole word feedback or phonological analysis feedback. This rank ordering was done because students began the study at staggered start times and we wanted to ensure that the full range of reading skills were represented in reach group. To prevent order effects words were randomized across student and session. This aspect of the experiment is a within subject design, where students act as their own control group (see Kirk, 1982). This experimental design is outlined in Figure 1.

**Feedback conditions.:** Students were randomly assigned to one of two feedback conditions. The first condition used a whole word (WW) feedback, which focused on presenting words as singular units. The second feedback technique for sight-word instruction, phoneme analysis (PhA) feedback method, provided the student with the word segmented into its sounds.

**Mastery.:** Mastery during word training was defined as 4 out of 5 consecutive exposures correct. This decision was made because students had four exposures to each word per day.

This criterion required students to succeed at either all exposures on one day or mastery across two days allowing for a single error.

**Procedure.** Examiners were graduate research assistants who had been trained on test administration until procedures were implemented with 90% fidelity. Fidelity was defined as the percentage of test procedures administered correctly. All tests were double-scored and double-entered; discrepancies were resolved by a third examiner. The average kappa across exposures approached 1.0 and agreement exceeded 99%. Average fidelity of test administration procedures (based on a random selection of 20% of the taped assessment sessions) exceeded 92% for all tests and word exposure trials.

**Data analysis.** Crossed-random effects models, a form of item-based analyses, were used to answer the research questions in this study (Van den Noortgate, De Boeck, & Meulders, 2003). These models were selected because they allowed us to partition the variance between child- and word-level predictors. Therefore, they are the only models that would allow us to simultaneously include child characteristics/skills (e.g., phonological awareness) and word characteristics (e.g., imageability) as predictors in the same models. A diagram of these item-based models is provided in Figure 2. For these models, words and persons are assumed to be random samples from a population of words and a population of persons. Since words are not nested within persons, these models are not strictly hierarchical models. Words and persons are on the same level and crossed in the design. Responses are nested within persons and within words. General equations for these models are provided in Appendix B. We conducted these analyses using Laplace approximation available through the lmer function (Bates & Maechler, 2009) from the lme4 library in R (R Development Team, 2012). The crossed-random effects models were built gradually in a stepwise fashion using model comparisons (likelihood ratio tests) to determine the model that best fit the data (Bates, 2011; Bates et al., 2015). We report word and child fixed effects coefficient estimates using gammas ( $\gamma$ ) and child-by-item fixed effect coefficients using lambdas ( $\lambda$ ). We use sigmas ( $\sigma$ ) to denote variance.

**Posttest model building.** The first model of interest addressed posttest word reading as the dependent variable. This unconditional model contained only random intercepts for child and word (i.e., no predictors) and was used as the base model for all subsequent model comparisons. This model was used to predict the probability of reading a target word correctly at posttest. The next model in the model building process included a child-by-word predictor of pretest target word reading. This item specific pretest performance predictor was included to control for prior knowledge of the words (i.e., the probability of reading the item correct at posttest conditioned on item performance at pretest). This model fit the data significantly better than the base model ( $\chi^2 = 39.38, p < 0.001$ ). A random slope for pretest target word reading was then added to the model, which did not fit the data significantly better than the previous model ( $\chi^2 = 6.96, p = .32$ ). Next, child- and word-level fixed effects were added to the model, which resulted in a significantly better fit ( $\chi^2 = 39.01, p < 0.001$ ). After this step, random slopes were added one at a time to determine which random slopes were needed for the best model fit. In the case of crossed-random effects models, a random slope refers to allowing a child characteristic to vary randomly across words and allowing a



word characteristic to vary randomly across children. A one-tailed significance test was used to test for random slopes. The only random slope that resulted in a significantly better model fit was a random slope for grade by word ( $\chi^2 = 6.292, p = .098$ ). All other model comparisons resulted in nonsignificant comparisons ( $p > .16$ ). In the posttest models, the WW group was the referent category and predicted probabilities are given for an average item and an average child in the WW group where all other covariates are at their mean values for our sample. For the final models that included a random slope in addition to the random intercepts, we calculated variance explained using the fixed slopes, random intercept models, a method supported by simulations by LaHuis, Hartman, Hakoyama, and Clark (2014).

**Maintenance model building.** Like the posttest model above, the first step in building the maintenance models was an unconditional model that included only random intercepts for child and word. In the next model, we included a child-by-word predictor of posttest performance (whether students read the word correctly at posttest). This model resulted in a significantly better fit than the base model ( $\chi^2 = 236.59, p < .001$ ). Fixed effects for child and word predictors were then added to the model. Random slopes were added individually for each of these predictors. There were no random slopes that resulted in a significantly better model fit ( $p > .27$ ).

**Mastery model building.** The process for building the mastery models was similar to that outlined for the other two models above. The dependent variable, however, was a continuous variable representing the number of exposures required in training for mastery. The unconditional model contained only a random intercept for child and word. The next model included a child-by-word predictor of posttest word reading performance (representing whether they read the word correctly at posttest). This model resulted in a significantly better fit ( $\chi^2 = 226.51, p < .001$ ). The next model included fixed effects for the predictor child and word features. Random slopes were then tested for each predictor. There were no other random slopes that resulted in a significantly better model fit ( $p > .55$ ).

## Results

Demographic data for the participants in this experiment ( $N=47$ ) are provided in Table 2. Table 3 provides child-level performance across measures disaggregated by condition with associated mean comparisons (ANOVAs). As noted in the table, there were no significant differences across groups on any of the pretest measures. Table 4 provides the zero order correlations amongst the child predictors of posttest word recognition. There were significant correlations between all child level predictors of word reading. The child level predictors were correlated with posttest reading performance with correlation absolute values ranging from .19 to .56.

There were three outcomes of interest in this experiment. The first was the probability of a correct response on a posttest measure of target word reading, the second was the probability of a correct response on the target word reading one week after posttest (maintenance), and the third was the number of exposures required for mastery of the words during training (mastery defined as 4 out of 5 consecutive attempts correct).

**Posttest target word reading.**—The unconditional model (not shown) for posttest performance (i.e., no child- or word-level predictors) that included only a random effect for child and word had an intercept of  $\gamma_{000} = 1.28$ , corresponding to a predicted probability of a correct response of .78 for the average child on the average word. Variability around that estimate was evident for both the child ( $= 5.63$ ) and word ( $= 1.81$ ). The final conditional models for posttest target word reading are presented in Table 5 (Models 1 & 2). The first model predicted posttest recognition of the target words, controlling for child-level pretest item performance (i.e., the probability of correctly reading the item *false* at posttest controlling for pretest performance reading the item), condition, PA, RAN, vocabulary, grade, and three planned comparisons for words (using low imageability irregular words as the referent group). Further execution of the model building process, as outlined above, resulted in a model that included a random intercept for word and child, a fixed effect for item-specific pretest, fixed effects for all child level and word level predictors, and a random slope for grade across words.

Significant predictors of posttest word reading were child-by-word pretest performance ( $\lambda_1=1.65$ ,  $z=5.16$ ,  $p<.001$ ); phonological awareness ( $\gamma_{004}=.23$ ,  $z=3.11$ ,  $p=.002$ ); rapid automatized naming ( $\gamma_{004}=-.02$ ,  $z=2.00$ ,  $p=.044$ ); word imageability ( $\gamma_{009}=1.44$ ,  $z=2.48$ ,  $p=.011$ ); and an interaction between regularity and imageability ( $\gamma_{010}=-1.71$ ,  $z=2.06$ ,  $p=.04$ ; see Model 1, Table 5). This interaction indicates that imageability appears to lead to a higher probability of reading target words correctly only for irregular words. In this model, child-level predictors of feedback condition, pretest vocabulary, and grade did not predict the probability of reading posttest items correctly. To further interpret these significant findings: these models indicate that an average child had a probability of .18 of reading a low imageability, irregular word correctly at posttest if she did not read it correctly at pretest. A student who did not read the word correctly at pretest but performed one standard deviation above the mean on PA and was average on all other measures had a probability of .34 of reading a low imageability irregular word correctly at posttest. A child with average scores on all child-level variables would have a probability of .48 of reading high imageability irregular words correctly.

**Maintenance of target word reading.**—The next model of interest addressed maintenance of word reading (see Model 2, Table 5). Controlling for pretest item performance, significant child-level predictors included PA ( $\gamma_{004}=.19$ ,  $z=2.62$ ,  $p=.002$ ) and grade ( $\gamma_{007}=1.53$ ,  $z=2.49$ ,  $p=.008$ ); significant word-level predictors included imageability ( $\gamma_{009}=.80$ ,  $z=2.03$ ,  $p=.04$ ) and the imageability by regularity interaction ( $\gamma_{008}=-1.32$ ,  $z=2.38$ ,  $p=.02$ ). Results indicate that when controlling for posttest item reading performance, the probability of a first-grade student reading a low imageability irregular word correctly at maintenance if they read the word correctly at pretest was .42. whereas the probability of a correct response was .04 if they did not read the word correctly at posttest. The final model from the model building process included a random effect for child and word and fixed effects for child-level and word-level predictors.

**Number of exposures required for mastery.**—The final model concerned the number of exposures required for children to reach mastery of the items (see Model 3, Table 6). All

students had 12 exposures to the words across three days. We defined mastery as four out of five exposures correct. In the mastery model, child-by-word posttest performance was controlled for because not all children achieved mastery on all items in 12 exposures. The unconditional model, which included only a random effect for word and child, indicated that an average of 8.28 exposures were required for mastery across items and children. In the model that controlled for posttest performance only, the number of exposures necessary for mastery was reduced to 7.31 if the word was read correctly at posttest. The mean number of exposures to mastery was 8.43 on low imageability irregular words. The number of exposures required to master low imageability regular words was 6.79, nearly two exposures less than low imageability irregular words.

The results from the first experiment indicated that significant predictors of posttest word reading performance were PA and RAN at the child level and imageability for irregular words at the word level. These findings suggest that students with good phonological skills are more likely to learn words and read words correctly at posttest. Furthermore, higher word imageability resulted in a higher probability of reading irregular words correctly after exposure. Likewise, students were more likely to maintain irregular words one week after posttest if they were highly imageable. Finally, students required fewer exposures to master low imageability regular words than they did to master low imageability irregular words. These findings suggest that both regularity and imageability impact both overall learning and rate of learning in students in first and second grade.

To further explore the role of imageability, we conducted a second experiment to address whether imageability could be used to enhance word reading instruction. To do so, we attempted to make low and high imageability words imageable for students by providing pictorial support prior to word reading instruction. We included a traditional vocabulary group and a word instruction only group as comparison groups.

## Experiment 2

### Method

**Participants.** For Experiment 2, 78 first grade children were drawn from one urban school district in the Southeastern region of the United States. Six first grade teachers were asked to send consent forms home with all students. We pretested all students using measures of word reading, pseudoword reading, phonological awareness, rapid automatized naming, and vocabulary. Students of all ability levels were included in the study.

### Child-level measures.

**Word reading.** The word reading task in this study was the Sight Word Efficiency task from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 2012). Students were asked to read a series of words in order of increasing difficulty for 45 seconds. The maximum score is 108 and the authors report an alternate forms reliability of .91.

**Pseudoword decoding.** To test pseudoword decoding skill, we used the Pseudoword Decoding Efficiency task from the TOWRE (Torgesen, Wagner, & Rashotte, 2012). Students were asked to read a list of pseudowords in order of increasing difficulty for 45 seconds. The maximum possible score is 66 and the authors report an alternate forms reliability of .92.

**Picture vocabulary.** The same vocabulary task was used in Experiment 1.

**Phonemic awareness (PA).** The phonemic awareness task was the Elision task from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, 2013). Students were asked to delete phonological units from words. The authors report test-retest reliability of .93.

**Rapid automatized naming (RAN).** To test for rapid automatized naming, we used the letter naming task from the CTOPP (Wagner, Torgesen, and Rashotte, 2013). For this task, students were asked to name a series of letters as fast as they could without making mistakes. The total score was the number of seconds students took to name all of the letters.

**Target word reading.** The target word reading task was a researcher developed task that required students to read a list of 32 words, which was the same list administered in Experiment 1. This list was scored in two steps, first for all 32 words, and then for only the 16 words targeted in the intervention (low imageability; see Table 7).

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

**Regularity.** Similar to Experiment 1, words were considered irregular if they were not consistent with typical letter sound correspondences. Regularity criteria was based on Rastle and Coltheart (1999).

**Experimental design.** The design of this experiment was similar to the design used in Experiment 1. Each student participated in the experiment for six consecutive days whenever possible, with everyone completing the experiment within 10 days and modifications being made to the experimental design when circumstances (e.g., attendance and school holidays) demanded it. Each group of students started the study on Friday and participated in a pretest battery of cognitive and reading related tests for approximately 30 minutes. Next, children were rank ordered based on their pretest scores on the target word reading test and then randomly assigned to one of three training conditions: (1) imageability (IMAG), (2) vocabulary (VOC), and (3) word-only (WO). As mentioned above, this rank ordering was done because students began the study at staggered start times and we wanted to ensure that the full range of reading skills were represented in each group. As in Experiment 1, to prevent order effects words were randomized across student and session. Following random assignment children participated in four days of training for approximately 15-20 minutes per day. Students then completed a short posttest battery for approximately 10-15 minutes. We only administered the target word reading task post-test. Similar to experiment 1 this was a within subject design, where students act as their own control group (see Kirk, 1982; see Figure 3).

**Conditions.:** On the first day of training, students in the imageability and the vocabulary conditions were exposed to either pictures (imageability) or verbal definitions for the words (vocabulary). Students in the word only condition read storybooks with a research assistant during that time. On the three remaining training days, students in the IMAG group spent 5 minutes reviewing pictures, students in the VOC condition spent 5 minutes reviewing definitions, and students in the WO condition spent 5 minutes reading storybooks before being exposed to each target word 4 times each day. To prevent order effects words were randomized across student and session. The number of times students heard the words was held constant across the imageability and vocabulary conditions. Corrective feedback was given at the whole word level across conditions.

**Imageability (IMAG).**—Materials for the imageability condition were created with a children’s illustrator. The illustrator helped to make the low imageability words imageable by creating a set of the pictures for the training. On the first day of training, students in the imageability condition were exposed to each picture twice and were encouraged to think of the picture when they heard that word. They were then asked to identify each picture when provided the word and asked to produce the word when presented with a picture. These activities were done using an imageability flipbook and game board. Using the game board, receptive and expressive activities were done at the beginning of the word learning sessions on days 2-4 of training. During word learning, students were encouraged to think of the picture for each word. The prompts used in the IMAG training are provided in Appendix C.

**Vocabulary (VOC).**—On the first day of training, students in the vocabulary condition were exposed to each definition twice and were encouraged to think of the meaning when they heard that word. Students were then asked to do an oral multiple choice task in which they were asked to identify each meaning when presented with the word and asked to produce the word when presented with the meaning. The multiple choice receptive and expressive vocabulary activities were done at the beginning of the word learning sessions on days 2-4 of training. During word learning, students were encouraged to think of the meaning for each word. The prompts used in the VOC training are provided in Appendix C.

**Word only (WO).**—In the word only condition, students read storybooks for 15 minutes on the first day and then for 5 minutes at the beginning of each word learning session to equalize instructional time. These storybooks did not include the target words. During word learning, students were only given corrective feedback at the whole word level, with no encouragement to think of the meaning or a picture for the words.

**Procedure.:** Test examiners were graduate research assistants trained on tests until procedures were implemented with 90% fidelity. All tests were double-scored and double-entered; discrepancies were resolved by a third examiner. 20% of exposure data was double entered. The average kappa across exposures was approaching 1.0 and agreement exceeded 99%. Average fidelity of procedures (based on a random selection of 20% of the taped sessions) exceeded 94% for all tests and training sessions.

**Data analysis.:** The analytic technique used in Experiment 2 was similar to the technique used in Experiment 1 (see above). The referent group for these models was the imageability

training group. Using IMAG as the referent group allowed us to compare IMAG vs. WO and IMAG vs. VOC. We include interaction terms in the mastery models to explore differences across training (with imageability as the referent group) based on initial reading skill (measured by the TOWRE). Thus, the interactions included were training group by TOWRE interactions.

## Results

Demographic data for the participants in this study ( $N=78$ ) are provided in Table 8. Table 9 provides child-level performance across measures disaggregated by condition with associated mean comparisons. As noted in the table, there were no significant differences across groups on any of the pretest measures. Table 10 provides the zero order correlations amongst the child level predictors of posttest word recognition. The child level predictors were correlated with posttest reading performance with correlation absolute values ranging from .11 to .69.

**Posttest target word reading.**—The unconditional model (not shown) for posttest performance included only a random effect for child and word and had an intercept of  $\gamma_{000} = 3.43$ , corresponding to a predicted probability of a correct response of .97 for the average child on the average word. Variability around that estimate was evident for both the child ( $= 10.38$ ) and word ( $= 1.71$ ). Further execution of the model building process (outlined in Experiment 1) resulted in a model that included a random effect for word and child, and fixed effects for all child level and word level predictors. The fixed effects for this model indicate that significant predictors of posttest performance were PA ( $\gamma_{005}=.22, z=3.43$ ), RAN ( $\gamma_{006}=-.08, z=3.08$ ), and word regularity ( $\gamma_{008}=1.66, z=3.15$ ; see Model 4, Table 11). These effects correspond to a probability of .93 for an average child reading a low imageability irregular word correctly at posttest if she did not read it correctly at pretest and .96 for students who read them correctly at pretest. Furthermore, there was a significant main effect for regularity. This indicates that a student with average scores on all child-level predictors who did not read the word correctly at pretest had an increased probability of .99 of reading a regular low imageability word at posttest, compared to a .93 probability for low imageability irregular words. Finally, there was no advantage to being in the IMAG condition over either the WO or VOC conditions in terms of posttest target word reading after controlling for child- and word-level predictors.

**Number of exposures required for mastery.**—These models are presented in Table 12. The unconditional model to address this question (not shown) contained only random effects for child and word. The intercept for this model was  $\gamma_{000} = 5.95$ , representing the average number of exposures required for mastery for the average child on the average word. Variability around that estimate was evident for both the child ( $= 2.15$ ) and word ( $= 0.72$ ). The model building process resulted in final models that included a random effect for word and child, a random slope for PA across words, and a random slope for TOWRE across words. Significant main effects (see Model 5, Table 12) include posttest item-specific target word reading, TOWRE Sight Word Efficiency ( $\gamma_{002}=-.07, z=4.98$ ), and the WO vs. imageability condition comparison ( $\gamma_{003}=.72, z=2.50$ ). The average number of exposures required for mastery of low imageability irregular words for students in the imageability

condition was 5.36 if students read the word correctly at posttest. The interaction model (see Model 6, Table 12) indicated that there was a significant interaction between condition and initial word reading skill on the TOWRE task ( $\gamma_{009} = -.04$ ,  $z = 2.24$ ). Students who started the intervention with low word reading skills benefited more from the imageability training than the word only training. This interaction is illustrated in Figure 4.

In sum, significant predictors of posttest word reading in experiment 2 were PA and RAN at the child level and regularity at the word level. Furthermore, there was a main effect for condition in efficiency of learning, with students in the imageability training group outperforming students in the word-only group. Finally, there was a significant interaction between imageability training and initial general word reading for word learning efficiency, with students starting with low word reading skills requiring fewer exposures for mastery in the imageability training group than in the word-only condition.

## Discussion

Results from these experiments both confirm and extend previous findings from the literature on the role of imageability in word reading. We found that imageability was an important factor in establishing a lexical representation of target words. These findings are consistent with others who have found that imageability ratings impact word learning differently across exposures (Duff & Hulme, 2012; Laing & Hulme, 1999) and the results from Experiment 2 offer preliminary evidence that imageability may be susceptible to instruction.

### The Role of Imageability in Word Reading Development

The first important finding from these experiments is that imageability seems to play a role in the development of word specific representations. Results from both studies suggest that imageability impacts the strength of a word representation across exposures. The findings for the posttest measure of word reading in the first experiment indicated that imageability was particularly important for irregular words. These results are similar to those of Duff and Hulme (2012), who found that imageability did not matter for earlier trials but did matter for later trials, with low imageability irregular words being the hardest to learn and the low imageability regular words being harder to learn than high imageability regular words. Our finding is also in line with the work of Strain, Patterson, and Seidenberg (1995) who found that imageability facilitates recognition of low-frequency irregular words and the findings of Strain and Herdman (1999), who found that reaction time for reading low imageability irregular words was much greater than reaction time for reading high imageability irregular words in adults.

According to Perfetti, “the major essential development in learning to read is the acquisition of individual word representations” (p.154) and “a word, once acquired, may be represented strictly as a specific unit” (p. 155). The significant role of imageability in these experiments could be interpreted as supporting children (particularly those with poor word reading skills) in creating these specific word representations. Although we did not find a main effect for the imageability training for posttest word reading in the second experiment, we did find a main effect for the imageability training in predicting the number of word exposures

required for mastery. Perhaps the visual supports in the imageability training and/or hearing the phonological form of the word multiple times serve as added supports in the word learning process, allowing students to establish stronger representations for words with fewer exposures than students in the word only condition. However, given that we found no difference between the imageability and vocabulary conditions, we can not conclude that the pictures are more effective than the definitions and/or hearing the phonological forms of the words. Further studies are required that specifically explore these contrasts. According to Perfetti's theory, we would expect that after several exposures, students would no longer require such supports once they have created a word specific representation. It is possible that 12 exposures were adequate for students to create a word specific representation, thereby resulting in no difference between the groups at posttest.

**The role of imageability in learning efficiency.**—A focus for both experiments was the efficiency of word learning and the role of child and word characteristics in the number of exposures children required for mastery. Using a criterion of four out of five exposures correct, we measured the number of exposures children required to master words. As we noted earlier, Ehri (1995) has demonstrated in the past that students at-risk for reading disabilities require more exposures to master words. Though our criterion for mastery was different than that for experiments done by Reitsma (1983) and Ehri (1995), our findings are in line with their general findings. Our studies extend previous work by exploring both child and word characteristics related to this continuous outcome. We found several factors that contribute to mastery at both the child- and word-levels. We found in the first experiment that when we controlled for pretest item-specific performance, low imageability regular words were easier to master than low imageability irregular words. In our second study when we intervened for imageability, we found an effect of condition, with students receiving imageability training requiring fewer exposures for mastery than students who were in the word only condition.

### Lexical Involvement in Word Learning

The results of our study and others suggest that imageability is particularly important for early learning of irregular words. Irregular words present unique challenges to students because they cannot rely exclusively on their knowledge of decoding rules to access lexical representations of irregular words. The results of our study seem to support others that associate lexical knowledge (e.g. of semantics or word forms) with learning to read irregular words. The importance of lexical knowledge has been demonstrated from several areas of the literature. The significant role of imageability in these two studies is consistent with the findings from studies on connectionist models of word recognition (e.g. Harm & Seidenberg, 2004; Plaut et al., 1996), which have been more successful at reading irregular words when a semantic processor is included in the model in addition to the orthographic and phonological processors. These findings are also consistent with studies that have found that either item-specific vocabulary knowledge (Ricketts, Nation, & Bishop, 2007) or familiarity with lexical phonology (McKague, Pratt, & Johnston, 2001; Nation & Cocksey, 2009; Taylor et al., 2011) are good predictors of reading irregular words correctly. These results may provide further support for Keenan and Betjemann's (2007) speculations that item-specific semantic activation may help to "fill voids" in phonological-orthographic processing in individuals



with poor mappings, such as children with reading difficulties (p. 193). Our results seem to support a model in which orthographic-to-phonological pathways become at least partially dependent on lexical input (e.g., from semantics), with this influence being increasingly important for irregular words (Nation & Snowling, 1998; Ricketts et al., 2007; Turner & Chapman, 2012).

### Imageability Training

Our attempt to increase imageability through training in Experiment 2 is one of several attempts to address imageability experimentally. As mentioned above, Duff and Hulme (2012) used nonwords to manipulate both phonological and semantic knowledge of their target words. They compared two conditions, one with only phonological information for the nonwords and one with both phonological and semantic information. They found no added benefit for semantic knowledge over and above the benefit of phonological information. Similarly, in our second experiment, we did not find a significant difference between the vocabulary and imageability groups for posttest performance. As noted earlier, the number of times students heard the words was equalized across these two conditions. It is possible that hearing the phonological representation was enough to reduce the number of exposures required for mastery. We speculate that the overall pool of words students had to choose from was reduced by the pretraining in both imageability and vocabulary, thus helping students to identify the words with fewer exposures. These findings are consistent with the findings of Wang, Nickels, Nation, and Castles (2013), who found that item-specific vocabulary knowledge was a predictor of orthographic learning only for irregular words. Although they did not focus only on the imageability, they trained on vocabulary using visual supports and thus have the same issue separating the benefits of vocabulary and imageability training as we do. It is clear that there remains a question within the literature regarding whether students only require phonological experience with words or whether visual supports and/or vocabulary definitions contribute something unique to the word learning process. Furthermore, the fact that we found an interaction favoring the imageability group for poor readers in the number of exposures required for mastery suggests that imageability supports may be beneficial only for our poorest readers. Further exploration of these questions is required.

### Limitations and Future Directions

The results from this study suggest that lexical input plays a role in overall word learning and efficiency of learning. The results also suggest that imageability is a potentially malleable word feature for instruction. There were several limitations, however, that should be considered when interpreting these results. First, this study focuses on only 32 words sampled from a large corpus of regular and irregular words. We are uncertain about how representative these words are of the entire corpus and thus we should exercise caution when generalizing these results to other words. Likewise, the sampling of children for both the at-risk sample and the representative sample was limited to the scope of the study. Both samples were relatively small and the results should be interpreted with this in mind. From an instructional perspective, there are several areas of interest for future studies. First, future studies could include a condition that pairs pictures with words rather than doing imageability training prior to word exposures. The experimental design and comparisons in

the present study did not allow us to determine the effects of such training but it may be beneficial for instruction. Next, a study exploring other feedback conditions is warranted. The feedback conditions in Experiment 1 may not have been the best forms of feedback and studies examining alternative feedback methods may be helpful. Finally, a study that examines irregular word reading instruction in more detail is warranted. For example, a study that encourages students to decode irregular words and look more carefully at the words at a subword level could be informative for both instruction and theory.

## Acknowledgments

We thank Leilani Dela Cruz for her artistic work on the illustrations for the imageability training. We also thank Alyson Collins, Johnny Daniel, Erika Hsu, Michael Jackson, Esther Lindstrom, and Meg Schiller for their help in data collection.

This research was supported in part by Grant P20HD091013 from NICHD. The content is solely the responsibility of the authors and does not necessarily represent the official view of NICHD.

## Appendix A

**Table 1A**

Descriptive Statistics for Target Word Characteristics

Word	Frequency (SFI)	Number of Letters	Orthographic Neighborhood Size	Phonological Neighborhood Size	Bigram Frequency by Position	Imageability	Age of Acquisition
<b>Irregular/High Imageability</b>							
bowl	55.2	4	9	1.35	563	579	256
laugh	55.6	5	0	1.95	709	528	230
soup	53	4	4	1.7	855	604	232
wolf	55.1	4	2	1.9	691	610	316
wound	53.4	5	8	1.55	1229	570	372
guard	56.2	5	0	1.95	801	530	344
world	68.4	5	1	1.85	1052	560	340
young	66.3	5	0	2	996	521	286
<b>Mean (SD)</b>	57.90 (5.96)	4.63 (.52)	3.00 (3.66)	1.78 (.23)	862.00 (218.62)	562.75 (34.40)	297.00 (54.20)
<b>Regular/High Imageability</b>							
birth	56.9	5	4	1.75	1069	532	328
hunt	56.1	4	8	1.35	877	527	366
farm	61.4	4	6	1.55	1478	560	306.5
ground	64.2	6	1	1.8	1604	513	258
brain	59	5	6	1.5	1914	572	347
coast	59.8	5	3	1.65	2601	588	311
space	63.7	5	4	1.6	1069	538	319
wife	60.9	4	9	1.55	465	575	314
<b>Mean (SD)</b>	60.25 (2.91)	4.75 (.71)	5.13 (2.64)	1.59 (.14)	1384.63 (667.54)	550.63 (26.77)	318.69 (31.71)
<b>Irregular/Low Imageability</b>							

Word	Frequency (SFI)	Number of Letters	Orthographic Neighborhood Size	Phonological Neighborhood Size	Bigram Frequency by Position	Imageability	Age of Acquisition
false	53.4	5	0	2	1245	315	399
build	61.5	5	2	1.85	779	399	251
broad	57.3	5	2	1.7	1348	463	376
learn	64.6	5	1	1.8	1512	361	280
lose	58.2	4	14	1.05	800	373	286
none	59	4	13	1.3	1656	425	267
sure	65.2	4	6	1.6	1547	285	328
worse	57.1	5	4	1.7	1501	339	357
<b>Mean (SD)</b>	59.54 (4.00)	4.63 (.52)	5.25 (5.42)	1.63 (.31)	1298.50 (338.03)	370.00 (58.33)	318.00 (54.91)
<b>Regular/Low Imageability</b>							
lie	57.5	3	9	1.05	426	385	250
choice	58.1	6	0	1.95	1328	303	379
cost	59.7	4	9	1.4	2731	401	417.5
plus	53.5	4	2	1.75	621	378	317
real	63.1	4	14	1.15	2432	313	310
stuff	54.7	5	4	1.75	941	305	372
trust	54.3	5	3	1.6	1287	356	400
which	73.3	5	0	2	1024	257	301
<b>Mean (SD)</b>	59.28 (6.49)	4.50 (.93)	5.13 (5.03)	1.58 (.35)	1348.75 (822.98)	337.25 (50.14)	343.31 (57.46)

## Appendix B

**Table 1B**

Crossed-Random Effects Models Used in Experiments 1 & 2

Base Models	
Posttest	Level 1 (Responses <sub>ijk</sub> ): $Logit(\pi_{ijk}) = \lambda_{0ji}$ Level 2 (Person <sub>j</sub> & Word <sub>i</sub> ): $\lambda_{0jk} = \gamma_{00} + r_{01j} + r_{02i}$ $r_{01j} \sim N(0, \sigma^2_{.01})$ & $r_{02i} \sim N(0, \sigma^2_{.02})$
Mastery	Level 1 (Responses <sub>ij</sub> ): $Y_{ij} = \lambda_{0jk} + e_{ijk}$ Level 2 (Person <sub>j</sub> & Word <sub>i</sub> ): $\lambda_{0jk} = \gamma_{00} + r_{01j} + r_{02i}$ $r_{01j} \sim N(0, \sigma^2_{.01})$ & $r_{02i} \sim N(0, \sigma^2_{.02})$ & $e_{ijk} \sim N(0, \sigma^2_e)$
Main Effects Models	
Posttest, Maintenance	Level 1 (Responses <sub>ij</sub> ): $Logit(\pi_{ji}) = \lambda_{0jk} + \sum_{c=1}^C \gamma_c^P c_{ij}$ Level 2 (Person <sub>j</sub> & Word <sub>i</sub> ):

$$\lambda_{0jk} = \gamma_{00} + \sum_{a=1}^A \gamma_a^M a_i + \sum_{b=1}^B \gamma_b^N b_j + r_{01j} + r_{02i}$$

$$r_{01j} \sim N(0, \sigma^2_{r01}) \& r_{02i} \sim N(0, \sigma^2_{r02})$$

Mastery Level 1 (Responses<sub>j</sub>):

$$Y_{ij} = \lambda_{0jk} + \sum_{c=1}^C \gamma_c^P c_{ij} + e_{ijk}$$

Level 2 (Person<sub>j</sub> & Word<sub>k</sub>):

$$\lambda_{0jk} = \gamma_{00} + \sum_{a=1}^A \gamma_a^M a_i + \sum_{b=1}^B \gamma_b^N b_j + r_{01j} + r_{02i}$$

$$r_{01j} \sim N(0, \sigma^2_{r01}) \& r_{02i} \sim N(0, \sigma^2_{r02}) \& e_{ijk} \sim N(0, \sigma^2_e)$$

Interaction Models

Mastery Level 1 (Responses<sub>j</sub>):

$$Y_{ij} = \lambda_{0jk} + \sum_{c=1}^C \gamma_c^P c_{ij} + e_{ijk}$$










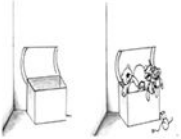




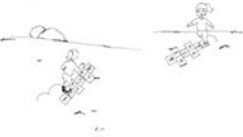

Level 2 (Person<sub>j</sub> & Word<sub>k</sub>):

$$\lambda_{0jk} = \gamma_{00} + \sum_{a=1}^A \gamma_a^M a_i + \sum_{b=1}^B \gamma_b^N b_j$$

$$+ r_{01j} + r_{02i} + \gamma_d^M 1_i^M 2_i + \gamma_d^M 1_i^M 3_i$$

$$r_{01j} \sim N(0, \sigma^2_{r01}) \& r_{02i} \sim N(0, \sigma^2_{r02}) \& e_{ijk} \sim N(0, \sigma^2_e)$$

**Appendix C**

<p style="text-align: center;">broad</p> 	<p style="text-align: center;">build</p> 	<p style="text-align: center;">cost</p> 	<p style="text-align: center;">learn</p> 
<p style="text-align: center;">cost</p> 	<p style="text-align: center;">went</p> 	<p style="text-align: center;">lie</p> 	<p style="text-align: center;">plus</p> 
<p style="text-align: center;">lose</p> 	<p style="text-align: center;">stuff</p> 	<p style="text-align: center;">false</p> 	<p style="text-align: center;">none</p> 
<p style="text-align: center;">trust</p> 	<p style="text-align: center;">choice</p> 	<p style="text-align: center;">once</p> 	<p style="text-align: center;">real</p> 

**Figure 1C.**  
Imageability training pictures.

<i>The word is...</i>	<i>It means...</i>	Read sentence	<i>When you hear the word...</i>
<b>went</b>	to leave	Sally <b>went</b> to the store.	<b>went</b> , remember that <b>went</b> means to leave.
<b>trust</b>	to have faith in	I <b>trust</b> my friends.	<b>trust</b> , remember that <b>trust</b> means to have faith in.
<b>plus</b>	adding together	I added two <b>plus</b> two.	<b>plus</b> , remember that <b>plus</b> means adding together.
<b>worse</b>	more bad	I feel <b>worse</b> than I did yesterday.	<b>worse</b> , remember that <b>worse</b> means more bad.
<b>learn</b>	to find out about something	We have so many things to <b>learn</b> at school.	<b>learn</b> , remember the word <b>learn</b> means to find out about something.
<b>cost</b>	the amount paid for something	Before you buy something, you find out the <b>cost</b> .	<b>cost</b> , remember the word <b>cost</b> means the amount paid for something.
<b>choice</b>	the act of choosing	You can have a <b>choice</b> between two things.	<b>choice</b> , remember the word <b>choice</b> means the act of choosing.
<b>stuff</b>	things that people need or use	We have lots of <b>stuff</b> in our garage.	<b>stuff</b> , remember that <b>stuff</b> means things that people need or use.
<b>lie</b>	to say something that is not true	I was taught to never tell a <b>lie</b> .	<b>lie</b> , remember that <b>lie</b> means to say something that is not true.
<b>none</b>	not any	I had lots of bananas and now I have <b>none</b> .	<b>none</b> , remember that <b>none</b> means not any.
<b>broad</b>	stretching far and wide	The eagle's wings are <b>broad</b> .	<b>broad</b> , remember the word <b>broad</b> means stretching far and wide.
<b>false</b>	not true	If something is untrue, it is <b>false</b> .	<b>false</b> , remember the word <b>false</b> means not true.
<b>lose</b>	to not be able to find	I hope I don't <b>lose</b> my favorite toy.	<b>lose</b> , remember that <b>lose</b> means to not be able to find.
<b>build</b>	to put together	Birds <b>build</b> nests for their homes.	<b>build</b> , remember that <b>build</b> means to put together.
<b>once</b>	one time only	Dinosaurs were <b>once</b> the largest creatures living on Earth.	<b>once</b> , remember that <b>once</b> means one time only.
<b>real</b>	not imaginary	I saw a <b>real</b> panda at the zoo.	<b>real</b> , remember that <b>real</b> means not imaginary.

**Figure 2C.**  
Vocabulary training definitions

## References

- Adams MJ (1994). *Beginning to read: Thinking and learning about print* The MIT Press, Cambridge, MA.
- Balota DA, Cortese MJ, Sergent-Marshall SD, Spieler DH, & Yap MJ (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133(2), 283. [PubMed: 15149254]
- Balota DA, & Ferraro FR (1993). A dissociation of frequency and regularity effects in pronunciation performance across young-adults, older adults, and individuals with senile dementia of the Alzheimer-type. *Journal of Memory and Language*, 32, 573–592.
- Bates D (2011, 6 22). [r-sig-ME] Correlated random effects in lmer and false convergence. Message posted to <https://stat.ethz.ch/pipermail/r-sig-mixed-models/2010q2/003921.html>.
- Bates D, & Machler M (2009). Package 'lme4' (Version 0.999375–32): linear mixed-effects models using S4 classes. Available (April 2011) at <http://cran.r-project.org/web/packages/lme4/lme4.pdf>.
- Bates D, Machler M, Bolker B, and Walker S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67 (1), 1–48.
- Bird H, Franklin S, & Howard D (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, 33(1), 73–79.
- Brasseur-Hock I, Hock MF, Kieffer MJ, Biancarosa G, & Deshler DD (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, 21, 438–452.
- Brybaert M, Warriner AB, & Kuperman V (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904–911. [PubMed: 24142837]

- Coltheart M (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, 4(1), 5–17.
- Coltheart V, & Leahy J (1992). Children's and adults' reading of nonwords: Effects of regularity and consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 718–729.
- Coltheart V, Laxon VJ, & Keating C (1988). Effects of word imageability and age of acquisition on children's reading. *British Journal of Psychology*, 79, 1–12 .
- Davies R, Barbón A, & Cuetos F (2013). Lexical and semantic age-of-acquisition effects on word naming in Spanish. *Memory & cognition*, 41(2), 297–311. [PubMed: 23180310]
- Denckla MB, & Rudel RG (1976). Rapid 'automatized' naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471–479 . [PubMed: 995240]
- Duff FJ, & Hulme C (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, 16, 504–525 .
- Dyson H, Best W, Solity J, & Hulme C (2017). Training Mispronunciation Correction and Word Meanings Improves Children's Ability to Learn to Read Words. *Scientific Studies of Reading*, 21(5), 392–407.
- Ehri LC (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, 18, 116–125.
- Ehri LC (1997). Sight word learning in normal readers and dyslexics In Blachman B (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 163–189). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Ehri LC (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1), 5–21.
- Ehri LC, & Saltmarsh J (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing*, 7(3), 295–326.
- Elbro C, de Jong PF, Houter D, & Nielsen A (2012). From spelling pronunciation to lexical access: A second step in word decoding? *Scientific Studies of Reading*, 16, 341–359.
- Foorman BR, & Torgesen J (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice*, 16(4), 203–212.
- García JR, & Cain K (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74–111.
- Gilhooly KJ, & Logie RH (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior research methods & instrumentation*, 12(4), 395–427.
- Grainger J, & Ziegler J (2011). A dual-route approach to orthographic processing. *Frontiers in psychology*, 2, 54. [PubMed: 21716577]
- Griffiths YM, & Snowling MJ (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology*, 94(1), 34–43.
- Harm MW, & Seidenberg MS (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological review*, 111(3), 662. [PubMed: 15250780]
- Johnston AM, Barnes MA, & Desrochers A (2008). Reading comprehension: Developmental processes, individual differences, and interventions. *Canadian Psychology/Psychologie canadienne*, 49(2), 125–132.
- Kaushanskaya M, & Rehtzigel K (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic bulletin & review*, 19(5), 935–941. [PubMed: 22638985]
- Keenan JM, & Betjemann RS (2007). Comprehension of single words: The role of semantics in word identification and reading disability. *Single-word reading: Behavioral and biological perspectives*, 191–210.
- Kirk RE (1982). *Experimental design*. John Wiley & Sons, Inc..
- LaHuis DM, Hartman MJ, Hakoyama S, & Clark PC (2014). Explained variance measures for multilevel models. *Organizational Research Methods*, 17(4), 433–451.

- Laing E, & Hulme C (1999). Phonological and semantic processes influence beginning readers' ability to learn to read words. *Journal of Experimental Child Psychology*, 73, 183–207. [PubMed: 10357872]
- Martin-Chang SL, & Levy BA (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing*, 18, 343–376 .
- Martin-Chang S, & Levy B (2006). Word reading fluency: A transfer appropriate processing account of fluency transfer. *Reading & Writing*, 19, 517–542.
- Martin-Chang SL, Levy BA, & O'Neil S (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology*, 96, 37–56. [PubMed: 17034812]
- McKague M, Pratt C, & Johnston MB (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*, 80(3), 231–262. [PubMed: 11274984]
- Metsala JL, Stanovich KE, & Brown GD (1998). Regularity effects and the phonological deficit model of reading disabilities: A meta-analytic review. *Journal of Educational Psychology*, 90, 279–293.
- Nation K & Cocksey J (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology*, 103(3), 296–308. [PubMed: 19410262]
- Ouellette GP (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98, 554–566.
- Ouellette G, & Fraser JR (2009). What exactly is a yait anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, 104(2), 239–251. [PubMed: 19481221]
- Paivio A (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255–287.
- Paivio A, Yuille JC, & Madigan SA (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of experimental psychology*, 76, 1–25.
- Perfetti CA (1992). The representation problems in reading acquisition In Gough PB, Ehri LC, & Treiman R (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.
- Perfetti C (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, 11(4), 357–383.
- Perfetti C, & Stafura J (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37.
- Perry C, Ziegler JC, & Zorzi M (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114(2), 273–315. [PubMed: 17500628]
- Pexman PM (2012). Meaning-based influences on visual word recognition. *Visual word recognition*, 2, 24–43.
- Plaut DC (1998). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science*, 23, 543–568.
- Plaut DC, McClelland JL, Seidenberg MS, & Patterson K (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological review*, 103(1), 56. [PubMed: 8650300]
- Plaut DC, & Shallice T (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive neuropsychology*, 10(5), 377–500.
- Priebe SJ, Keenan JM, & Miller AC (2012). How prior knowledge affects word identification and comprehension. *Reading and writing*, 25, 131–149.
- Pritchard SC, Coltheart M, Marinus E, & Castles A (2018). A Computational Model of the Self-Teaching Hypothesis Based on the Dual-Route Cascaded Model of Reading. *Cognitive science*.
- R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria ISBN 3–900051-07–0, URL <http://www.R-project.org/>.
- Rastle K, & Coltheart M (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human perception and performance*, 25(2), 482.



- Rayner K, Foorman BR, Perfetti CA, Pesetsky D, & Seidenberg MS (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2), 31–74. [PubMed: 26151366]
- Reitsma P (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, 36, 321–339.
- Ricketts J, Davies R, Masterson J, Stuart M, & Duff FJ (2016). Evidence for semantic involvement in regular and exception word reading in emergent readers of English. *Journal of experimental child psychology*, 150, 330–345. [PubMed: 27416563]
- Ricketts J, Nation K, & Bishop DVM (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11(3), 235–257.
- Rosner J (1979). Test of Auditory Analysis Skills (TAAS) In: *Helping children overcome learning difficulties: A step-by-step guide for parents and teachers* (pp. 77–80). New York: Academic Therapy.
- Sabsevitz DS, Medler DA, Seidenberg M, & Binder JR (2005). Modulation of the semantic system by word imageability. *Neuroimage*, 27(1), 188–200. [PubMed: 15893940]
- Seidenberg MS, & McClelland JL (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523–568. [PubMed: 2798649]
- Seidenberg MS, Waters GS, Barnes MA, & Tanenhaus MK (1984). When does irregular spelling or pronunciation influence word recognition?. *Journal of Verbal Learning and Verbal Behavior*, 23, 383–404.
- Seymour PH, Aro M, & Erskine JM (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174. [PubMed: 12803812]
- Spaai GG, Ellermann HH, & Reitsma P (1991). Effects of segmented and whole-word sound feedback on learning to read single words. *Journal of Educational Research*, 84, 204–213.
- Stanovich KE (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading research quarterly*, 32–71.
- Stanovich KE, & Bauer DW (1978). Experiments on the spelling-to-sound regularity effect in word recognition. *Memory & Cognition*, 6, 410–415.
- Steady LM, Kearns DM, Gilbert JK, Compton DL, Cho E, Lindstrom ER, & Collins AA (2017). Exploring individual differences in irregular word recognition among children with early-emerging and late-emerging word reading difficulty. *Journal of Educational Psychology*, 109(1), 51–69.
- Strain E, Patterson K, & Seidenberg MS (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(5), 1140.
- Strain E, & Herdman CM (1999). Imageability effects in word naming: an individual differences analysis. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 53, 347–359. [PubMed: 10646206]
- Taylor JN, & Perfetti CA (2016). Eye movements reveal readers' lexical quality and reading experience. *Reading and Writing*, 29(6), 1069–1103.
- Taylor JSH, Plunkett K, & Nation K (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 60.
- Toglia MP, & Battig WF (1978). *Handbook of semantic word norms*. Lawrence Erlbaum.
- Torgesen JK, Wagner R, & Rashotte C (2012). *Test of Word Reading Efficiency 2*. Austin, TX: Pro-Ed.
- van Daal VP, & Reitsma P (1990). Effects of independent word practice with segmented and whole-word sound feedback in disabled readers. *Journal of Research in Reading*, 13, 133–48.
- Van den Noortgate W, De Boeck P, & Meulders M (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics*, 28(4), 369–386.
- Wagner RK, Torgesen JK, Rashotte C, & Pearson NA (2013). *Comprehensive Test of Phonological Processing 2*. Austin, TX: Pro-Ed.
- Wang HC, Nickels L, Nation K, & Castles A (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, 17, 369–384.
- Waters GS, & Seidenberg MS (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition*, 13, 557–572.

Woodcock RW (2011). Woodcock reading mastery tests: WRMT-III. Manual. Pearson.

Woodcock RW, McGrew KS, & Mather N (2001). Woodcock-Johnson III Tests of Achievement. Rolling Meadows, IL: Riverside Publishing.

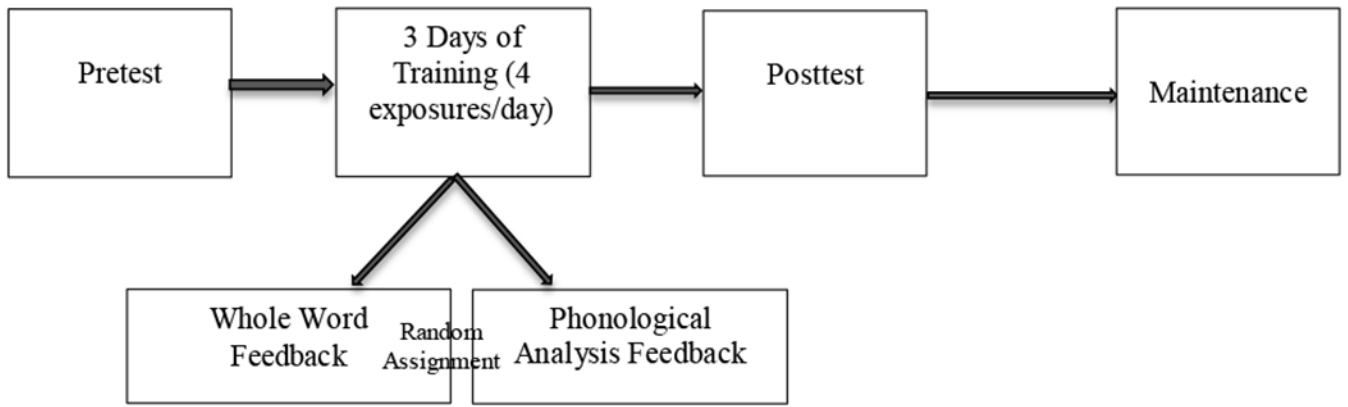
Ziegler JC, Perry C, & Zorzi M (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634).

Author Manuscript

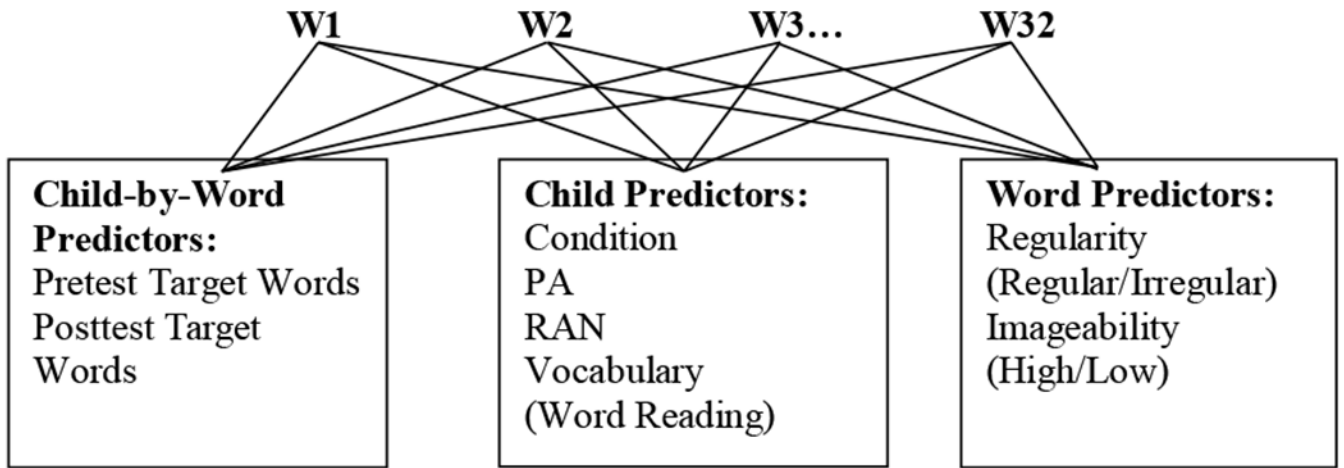
Author Manuscript

Author Manuscript

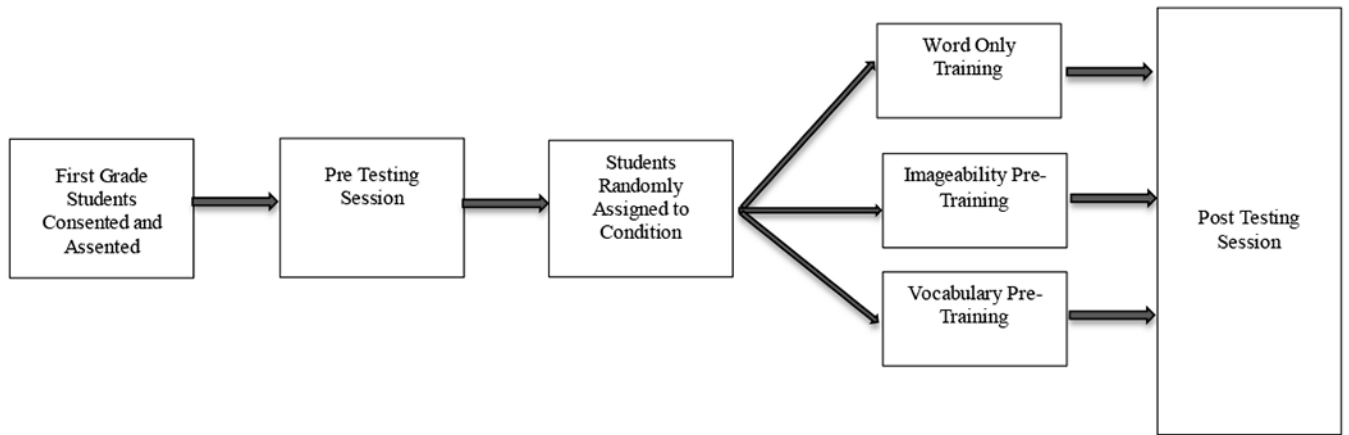
Author Manuscript



**Figure 1.**  
Experimental design for Experiment 1.



**Figure 2.**  
Crossed-random effects models for Experiment 1 and Experiment 2.



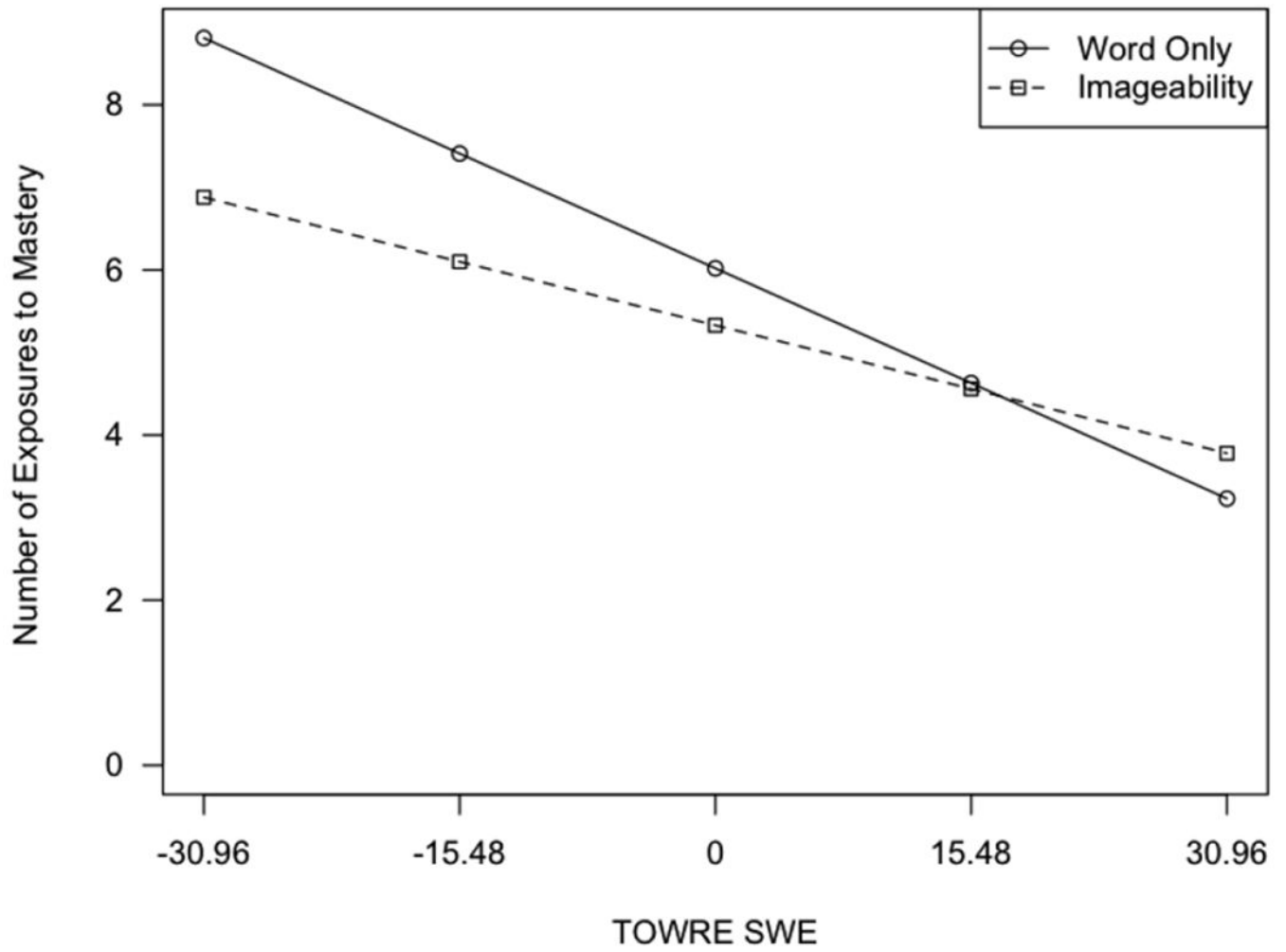
**Figure 3.**  
Experimental design for Experiment 2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 4.** Interaction between initial word reading (TOWRE Sight Word Efficiency) and condition in Experiment 2.

**Table 1**

Word Factorial Design for Experiment 1

	<b>Regular</b>	<b>Irregular</b>
<b>High Imageability</b>	brain	bowl
	birth	eye
	coast	foot
	farm	guard
	ground	laugh
	hunt	soup
	space	world
	wife	young
<b>Low Imageability</b>	choice	broad
	cost	build
	lie	false
	plus	learn
	real	lose
	stuff	none
	trust	once
	went	worse

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

## Demographic Statistics for Experiment 1

Variable	<i>N</i> = 47			
	<i>n</i>	%	Mean	(SD)
Age (years)			7.08	(.60)
Gender				
Male	28	59.57		
Female	19	40.43		
Grade				
1	29	61.70		
2	18	38.30		
Group				
Whole Word	25	53.19		
Phonological Analysis	22	46.81		
Race				
African American	3	6.38		
Hispanic	3	6.38		
Caucasian	40	85.12		
Biracial	1	2.12		

*Note:* Age was calculated based on age at the outset of the study.



**Table 3**

Child Level Descriptive Statistics for Experiment 1

Variable	<u>Whole Word</u>		<u>Phonological Analysis</u>		<u>All children</u>		<u>Pairwise comparisons<sup>a</sup></u>
	M	(SD)	M	(SD)	M	(SD)	
	<i>n</i> = 25		<i>n</i> = 22		<i>n</i> = 47		
Word Identification (RS)	10.92	(6.09)	10.95	(6.07)	10.94	(6.01)	WW= PhA
Word Identification (SS)	89.72	(12.45)	89.95	(11.63)	89.83	(11.94)	WW= PhA
Word Attack (RS)	3.76	(3.47)	4.59	(3.58)	4.15	(3.51)	WW= PhA
Word Attack (SS)	88.16	(11.64)	90.55	(10.97)	89.28	(11.27)	WW= PhA
PA	9.00	(3.74)	7.77	(3.96)	8.43	(3.85)	WW= PhA
RAN	42.00	(9.56)	51.09	(29.33)	46.26	(21.48)	WW= PhA
VOC (RS)	17.56	(2.86)	18.05	(3.50)	17.79	(3.15)	WW= PhA
VOC (SS)	92.88	(9.63)	94.54	(11.34)	93.66	(10.38)	WW= PhA

*Note:* RS=Raw Score; SS=Standard Score; PA = Phonological Awareness; RAN=Rapid Automatized Naming; VOC= Vocabulary; WW = Whole word feedback; PhA = Phonological Analysis Feedback.

<sup>a</sup>Mean comparisons were conducted using ANOVA with Bonferonni post-hoc pairwise comparisons.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**

Zero Order Correlations Between Child Variables: Experiment 1

	1	2	3	4	5
1 Pretest Word Reading	–				
2 PA	<b>.56</b>	–			
3 RAN	<b>–.46</b>	<b>–.37</b>	–		
4 VOC	<b>.40</b>	.22	–.19	–	
5 Posttest Word Reading	<b>.82</b>	<b>.58</b>	<b>–.51</b>	<b>.30</b>	–

Note:  $p < .001$  for all variables in bold. PA = Phonological Awareness; RAN=Rapid Automatized Naming; VOC= Vocabulary

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5**

Fixed Effects and Variance Estimates for Target Word Recognition for Experiment 1

Fixed Effects Parameter	Model 1			Model 2		
	Posttest Model			Maintenance Model		
	Est.	(SE)	z	Est.	(SE)	z
Intercept ( $\gamma_{000}$ )	-1.52	(.99)	-1.51	<b>-3.22</b>	<b>(.73)</b>	<b>4.42</b>
Item covariates						
$\lambda_1$ pretest	<b>1.65</b>	<b>(.32)</b>	<b>5.18</b>	—	—	—
$\lambda_2$ posttest	—	—	—	<b>2.90</b>	<b>.20</b>	<b>14.51</b>
Child covariates						
$\gamma_{003}$ Condition	.56	(.48)	1.16	.58	(.37)	1.54
$\gamma_{004}$ PA	<b>.23</b>	<b>(.07)</b>	<b>3.11</b>	<b>.12</b>	<b>(.06)</b>	<b>2.23</b>
$\gamma_{005}$ RAN	<b>-.02</b>	<b>(.01)</b>	<b>-2.00</b>	-.01	(.01)	-1.22
$\gamma_{006}$ Vocabulary	.06	(.08)	.71	.01	(.06)	.15
$\gamma_{007}$ Grade	1.02	(.65)	1.58	<b>1.18</b>	<b>(.48)</b>	<b>2.44</b>
Word covariates						
$\gamma_{008}$ Regularity	1.01	(.57)	1.77	.73	(.39)	1.86
$\gamma_{009}$ Imageability	<b>1.44</b>	<b>(.58)</b>	<b>2.48</b>	<b>.80</b>	<b>(.40)</b>	<b>2.03</b>
$\gamma_{010}$ Regularity*Imageability	<b>-1.71</b>	<b>(.83)</b>	<b>2.06</b>	<b>-1.32</b>	<b>(.56)</b>	<b>2.38</b>
	% var explained			% var explained		
Person	54.50			25.17		
Word	16.55			44.93		

**Table 6**

Experiment 1: Fixed Effects and Variance Estimates for Number of Exposures Required for Mastery

Fixed Effects Parameter	Mastery Model (Model 3)		
	Est.	(SE)	t
Intercept ( $\gamma_{000}$ )	<b>10.56</b>	<b>.47</b>	<b>22.59</b>
Item covariate			
$\lambda_1$ posttest	<b>-2.13</b>	<b>.15</b>	<b>-14.08</b>
Child covariates			
$\gamma_{002}$ Condition	-.52	.40	-1.30
$\gamma_{003}$ PA	<b>-.21</b>	<b>.06</b>	<b>-3.84</b>
$\gamma_{004}$ RAN	<b>.03</b>	<b>.01</b>	<b>3.07</b>
$\gamma_{005}$ Vocabulary	-.11	.06	-1.76
Word covariates			
$\gamma_{006}$ Regularity	<b>-1.64</b>	<b>.54</b>	<b>-3.06</b>
$\gamma_{007}$ Imageability	-.99	.54	-1.85
$\gamma_{008}$ Regularity*Imageability	<b>1.84</b>	<b>.76</b>	<b>2.42</b>
% var explained			
Person	50.00		
Word	17.69		

Note: PA = Phonological awareness (elision); RAN = Rapid automatic naming; Imag. = Imageability

$p < .05$  for variables in bold

**Table 7**

Word Lists for Experiment 2

	<b>Regular</b>	<b>Irregular</b>
<b>Low Imageability</b>	choice	Broad
	cost	build
	lie	false
	plus	learn
	real	lose
	stuff	none
	trust	once
	went	worse

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 8**

## Demographic Statistics for Experiment 2

Variable	<i>N</i> = 78		Mean	(SD)
	<i>n</i>	%		
Age (years)			7.08	(.60)
Gender				
Male	42	47.73		
Female	33	42.31		
Unreported	3	3.85		
Group				
Word Only	26	33.33		
Vocabulary	26	33.33		
Imageability	26	33.33		
Race				
African American	14	17.95		
Hispanic	8	10.26		
Caucasian	51	65.38		
Unreported	8	10.26		

*Note:* Age was calculated based on age at the outset of the study.

**Table 9**

Child Level Descriptive Statistics for Experiment 2

Variable	Word Only		Vocabulary		Imageability		All children		Pairwise comparisons <sup>a</sup>
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	
	<i>n</i> = 26		<i>n</i> = 26		<i>n</i> = 26		<i>n</i> = 78		
SWE (RS)	30.88	(14.28)	31.31	(16.31)	34.12	(16.44)	32.10	(15.57)	WO=VOC=IMAG
SWE (SS)	101.46	(14.32)	101.69	(16.21)	104.58	(16.14)	102.58	(15.45)	WO=VOC=IMAG
PDE (RS)	13.31	(8.54)	13.15	(8.76)	12.08	(8.48)	12.85	(8.50)	WO=VOC=IMAG
PDE (SS)	101.15	(16.82)	101.35	(16.36)	99.46	(15.92)	100.65	(16.17)	WO=VOC=IMAG
PA (RS)	16.54	(7.64)	18.31	(5.99)	17.42	(7.63)	17.42	(7.07)	WO=VOC=IMAG
PA (SS)	10.31	(2.85)	10.92	(2.46)	10.88	(3.05)	10.71	(2.78)	WO=VOC=IMAG
RAN (RS)	32.65	(11.34)	34.38	(13.21)	33.00	(12.31)	33.35	(12.17)	WO=VOC=IMAG
RAN (SS)	10.19	(1.30)	9.88	(1.77)	10.35	(1.50)	10.14	(1.53)	WO=VOC=IMAG
VOC (RS)	18.38	(3.72)	19.15	(2.31)	18.31	(3.06)	18.62	(3.07)	WO=VOC=IMAG
VOC (SS)	98.73	(13.12)	101.50	(7.75)	98.35	(10.68)	99.53	(10.70)	WO=VOC=IMAG

Note: SWE=TOWRE Sight Word Efficiency; PDE=TOWRE Pseudoword Decoding; PA = Phonological Awareness; RAN=Rapid Automatized Naming; VOC= Vocabulary

<sup>a</sup>Mean comparisons were conducted using ANOVA with Bonferonni post-hoc pairwise comparisons.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 10**

Zero Order Correlations between child variables: Experiment 2

	1	2	3	4	5
1 Pretest Word Reading	–				
2 PA	<b>.69</b>	–			
3 RAN	<b>–.43</b>	<b>–.35</b>	–		
4 VOC	<b>.37</b>	<b>.41</b>	–.11	–	
5 Posttest Word Reading	<b>.71</b>	<b>.45</b>	<b>–.48</b>	.19	–

Note:  $p < .001$  for all variables in bold. PA = Phonological Awareness; RAN=Rapid Automatized Naming; VOC= Vocabulary

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 11**

Fixed Effects and Variance Estimates for Target Word Recognition for Experiment 2

<b>Model 4</b>			
<b>Posttest Model</b>			
<b>Fixed Effects Parameter</b>	<b>Est.</b>	<b>(SE)</b>	<b>z</b>
Intercept ( $\gamma_{000}$ )	<b>2.56</b>	<b>(.67)</b>	<b>3.83</b>
Item covariates			
$\lambda_1$ pretest	.58	(.34)	1.69
$\lambda_2$ posttest	—	—	—
Child covariates			
$\gamma_{003}$ WO vs. IMAG Cond.	.17	(.78)	.21
$\gamma_{004}$ VOC vs. IMAG Cond.	-.89	(.78)	-1.14
$\gamma_{005}$ PA	<b>.22</b>	<b>(.06)</b>	<b>3.48</b>
$\gamma_{006}$ RAN	<b>-.08</b>	<b>(.03)</b>	<b>-3.08</b>
$\gamma_{007}$ Vocabulary	.06	(.12)	.47
Word covariate			
$\gamma_{008}$ Regularity	<b>1.66</b>	<b>(.53)</b>	<b>3.15</b>
% var explained			
Person	50.84		
Word	48.28		

*Note:* PA = Phonological awareness (elision); RAN = Rapid automatic naming; IMAG. = Imageability; VOC = Vocabulary; WO = Word Only; Cond.=Condition

$p < .05$  for variables in bold

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 12**

Experiment 2: Fixed Effects and Variance Estimates for Number of Exposures Required for Mastery.

Fixed Effects Parameter	Model 5			Model 6		
	Mastery Model			Interaction Model		
	Est.	(SE)	t	Est.	(SE)	t
Intercept ( $\gamma_{000}$ )	<b>7.51</b>	<b>(.29)</b>	<b>25.85</b>	<b>7.49</b>	<b>(.29)</b>	<b>26.02</b>
Item covariate						
$\lambda_1$ posttest	<b>-2.15</b>	<b>(.18)</b>	<b>-11.87</b>	<b>-2.16</b>	<b>(.18)</b>	<b>-11.98</b>
Child covariates						
$\gamma_{002}$ TOWRE	<b>-.07</b>	<b>(.01)</b>	<b>-4.98</b>	<b>-.05</b>	<b>(.02)</b>	<b>-3.22</b>
$\gamma_{003}$ WO vs. IMAG Condition	<b>.72</b>	<b>(.29)</b>	<b>2.50</b>	<b>.69</b>	<b>(.28)</b>	<b>2.44</b>
$\gamma_{004}$ VOC vs. IMAG Condition	.14	(.29)	.49	.16	(.28)	.58
$\gamma_{005}$ PA	-.04	(.02)	-1.79	-.04	(.02)	-1.86
$\gamma_{006}$ RAN	.01	(.01)	.98	.01	(.01)	.98
$\gamma_{007}$ Vocabulary	.00	(.04)	.00	.01	(.04)	.33
Word covariate						
$\gamma_{008}$ Regularity	-.22	(.15)	-1.51	-.22	(.15)	-1.51
Interactions						
$\gamma_{009}$ TOWRE*WO vs. IMAG	—	—	—	<b>-.04</b>	<b>(.02)</b>	<b>-2.24</b>
$\gamma_{010}$ TOWRE*VO vs. IMAG	—	—	—	-.01	(.02)	-.76
	% var explained			% var explained		
Person	59.66			61.00		
Item	27.81			27.86		

Note: PA = Phonemic awareness (elision); RAN = Rapid automatized naming; IMAG. = Imageability; VOC = Vocabulary; WO = Word Only.

$p < .05$  for variables in bold.