Learning Spellings and Meanings: Longitudinal Relations to Reading

Catherine Mimeau[1], Jessie Ricketts[2], and S. Hélène Deacon[1]

[1] Department of Psychology and Neuroscience, Dalhousie University

[2] Department of Psychology, Royal Holloway, University of London

Author Note

Catherine Mimeau https://orcid.org/0000-0001-6497-2392

Jessie Ricketts https://orcid.org/0000-0002-2760-3990

S. Hélène Deacon https://orcid.org/0000-0002-4792-5137

Catherine Mimeau is now at Département Sciences humaines, Lettres et Communication, Université TÉLUQ.

This study has been approved by the Social Sciences and Humanities Research Ethics Board of Dalhousie University. The data that support the findings of this study are available from the corresponding author upon reasonable request. Data from the first time point of this study (Grade 3) were presented in Mimeau et al. (2018). We have no conflicts of interest to disclose. This work was funded by a Postdoctoral Fellowship awarded to Catherine Mimeau by the Fonds de recherche du Québec – Société et culture, Grant ES/K008064/1 awarded to Jessie Ricketts by the UK Economic and Social Research Council, and Grant RGPN 293300-2013 awarded to S. Hélène Deacon by the Natural Sciences and Engineering Research Council of Canada. We are grateful to Ekaterina Reymarova, Laura Aziz, Hannah Robinson, Tammy Lamb, Carlee MacNeill, Elizabeth MacKay, and Jabran Umar for their help with testing and data scoring and entry.

Correspondence concerning this article should be addressed to Catherine Mimeau, Département Sciences humaines, Lettres et Communication, Université TÉLUQ, 455 rue du Parvis, Québec (Quebec) G1K 9H6, Canada. E-mail: catherine.mimeau@teluq.ca

Abstract

Background: Prominent theories of reading make the prediction that individual differences in children's word learning capacity determine the pace of their acquisition of reading skill. Despite the developmental nature of some of these theories, most empirical research to date has explored the relation between word learning capacity and reading at a single time point. The present study extends this research base by investigating whether earlier learning of the spelling and meaning of words is associated with later core aspects of reading: orthographic representations, word reading, and reading comprehension. **Methods**: Participants were 120 English-speaking children followed longitudinally from Grade 3 to Grade 4 (i.e., from 8 to 9 years of age on average). At Grade 3, children read stories containing new words and answered questions about the spelling and meaning of these new words, evaluating orthographic and semantic learning, respectively. Children also completed outcome measures of orthographic representations (with a choice task targeting the spelling of existing words), word reading, and reading comprehension (with standardised tasks) at Grades 3 and 4. We conducted regression analyses controlling for age, non-verbal reasoning, working memory, vocabulary, and phonological awareness. **Results**: We found that each of orthographic and semantic learning predicted gains in orthographic representations from Grade 3 to Grade 4. Furthermore, orthographic learning at Grade 3 predicted word reading at Grade 4, while semantic learning at Grade 3 predicted reading comprehension at Grade 4. **Conclusions**: These longitudinal associations between orthographic and semantic learning and core aspects of reading strengthen the evidence in support of the hypothesis that children's word learning capacity plays a key role in reading development.

Keywords: orthographic learning, semantic learning, reading development, longitudinal study

Implications for Practice

What is already known about this topic

- Individual differences in children's capacity to learn the spelling and meaning of new words during independent reading are associated with their mental representations of words' spelling, their ability to read words, and their ability to understand texts at the same point in time.
- Prominent theories of reading imply that individual differences in children's word learning capacity play a key role in the development of reading over time.

What this paper adds

- Individual differences in children's capacity to learn the spelling and the meaning of new words during independent reading can predict how much their mental representations of words' spelling will improve over one year.
- Individual differences in children's capacity to learn the spelling of new words during independent reading can predict how well they will read words one year later.
- Individual differences in children's capacity to learn the meaning of new words during independent reading can predict how well they will understand texts one year later.

Implications for theory, policy or practice

- The longitudinal associations between orthographic and semantic learning and core aspects of reading strengthen the evidence in support of the hypothesis that children's word learning capacity plays a key role in reading development.
- Our findings encourage reading programs that foster the development of cognitive and behavioral strategies to support word learning during independent reading.

Learning Spellings and Meanings: Longitudinal Relations to Reading

Reading is one of the most important skills that children develop during the elementary school years. In particular, understanding the meaning of texts, or reading comprehension, is the ultimate goal of reading development. Key theories of reading can explain the steps necessary to reach this goal. Foremost, in order to achieve reading comprehension, children need to develop the ability to read the individual words they encounter efficiently (simple view of reading; Gough & Tunmer, 1986). Efficient word reading skill, in turn, requires children to match the words they encounter with a mental representation of their spelling and to access stored information about their pronunciation (and meaning; lexical quality hypothesis; Perfetti & Hart, 2002). The formation of mental representations of words' spellings, that is, of orthographic representations, is deemed to depend on children's word learning capacity (self-teaching-hypothesis; Share, 1995). This latter idea is a noteworthy shift from the investigation of children's crystallised knowledge to that of their fluid abilities, effectively, from acquired knowledge to the capacity to learn. Given the developmental nature of the ideas conveyed by these prominent theories of reading, we used a longitudinal design to test them in the present study. Specifically, we investigated whether skill in learning the spelling and meaning of words could predict, one year later, three aspects of reading: orthographic representations, word reading, and reading comprehension.

Orthographic Learning and Reading

Once children can read, the reading process provides opportunities for word learning. For example, in his self-teaching hypothesis, Share (1995) postulates that children's letter by letter decoding enables their *orthographic learning*: the learning of the spelling of new words (see also Ehri's amalgamation theory, 2020, and the implementation of self-teaching within the dual-route cascaded model of reading, Pritchard et al., 2018). In a typical orthographic learning paradigm, children read stories containing non-words independently and are then asked to recall or recognise the spelling of these non-words. According to Share, orthographic learning allows children to create orthographic representations, that is, mental representations of the spelling of new words, so that children can recognise and read these words more efficiently in subsequent encounters. Consequently, Share's self-teaching hypothesis implies that early individual differences in children's orthographic learning are a predictor of later individual differences in their orthographic representations and word reading. Given that word reading is necessary (although not sufficient) for reading comprehension, as postulated by the simple view of reading (Gough & Tunmer, 1986), orthographic learning should also influence reading comprehension through its effect on word reading.

In line with these theoretical propositions, empirical studies have uncovered cross-sectional associations between children's orthographic learning and their orthographic representations, word reading, and reading comprehension (e.g., Deacon, Mimeau, et al., 2019; Mimeau et al., 2018; Ricketts et al., 2011). Further, in recent research, the association between orthographic learning and word reading was found to remain significant even after taking into account several control variables (Deacon, Mimeau, et al., 2019; Mimeau et al., 2018). This finding suggests a unique contribution of orthographic learning to word reading, as predicted by Share's (1995) self-teaching hypothesis. In contrast, the association between orthographic learning and reading comprehension was fully mediated by word reading (Mimeau et al., 2018), in line with the simple view of reading (Gough & Tunmer, 1986).

Building on these studies with a concurrent design, a core next question lies in testing the developmental nature of the relation between orthographic learning and reading. Longitudinal studies provide an important step towards answering this question. Indeed, although studies conducted at one point in time seem to suggest that orthographic learning predicts word reading, findings from these studies could just as well be interpreted in the opposite direction (e.g., Ricketts et al., 2011). This idea is in line with the well-established Matthew effects, according to which children who are good readers will get more out of their reading than children with poorer reading skill (Stanovich, 2009).

There are remarkably few available longitudinal studies on the relation between orthographic learning and reading. As an example, Deacon, Pasquarella, et al. (2019) followed a group of 8- and 9-year-old English-speaking children from Grades 2 and 3 (Time 1) to Grades 3 and 4 (Time 2) in a study on independent reading. The authors showed that orthographic learning at Time 1 predicted a latent factor combining orthographic representations and word reading at Time 2 (see Wolter et al., 2011, for similar findings in younger children engaging in shared book reading). The relation in Deacon, Pasquarella, et al.'s study held after controlling for non-verbal reasoning, phonological awareness, and orthographic representations and word reading at Time 1. However, because a latent factor was used, it remains unclear whether orthographic learning can predict each of orthographic representations and word reading separately.

Semantic Learning and Reading

Beyond new words' spellings, children are also exposed to new words' meanings when they read texts independently. The lexical quality hypothesis (Perfetti & Hart, 2002) and triangle models of reading (e.g., Seidenberg & McClelland, 1989) propose that this semantic information is crucial for reading development (see also Ehri, 2020). Indeed, these theories suggest that the development of efficient word reading, and thereby effective reading comprehension, depends on the quality of one's representations of words, including their semantic representations. Furthermore, Seidenberg and McClelland (1989) posit that all aspects of word representations—phonological, orthographic, and semantic—influence each other; the formation of semantic representations should thus influence the development of orthographic representations. Empirical research also points to a causal role of semantics in both word reading and reading comprehension (e.g., Quinn et al., 2015; Taylor et al., 2015).

Bringing this focus on semantics to self-teaching in reading, one could extend Share's (1995) hypothesis. The self-teaching hypothesis implies that individual differences in children's orthographic learning can predict individual differences in their orthographic representations and word reading. It can then be speculated that individual differences in children's acquisition of semantic representations, or *semantic learning*, could be another factor determining individual differences in children's reading (see Mimeau et al., 2018). For instance, if a child easily learns the meaning of new written words, their semantic representations of these words may be strengthened, facilitating the reading of these words. Reading comprehension could also benefit from the quick formation of high-quality semantic representations of words, considering the importance of words' meaning for understanding texts (Perfetti & Hart, 2002; Quinn et al., 2015).

Building on the extensive body of research on semantic learning during shared book reading (see Flack et al., 2018, for a review), newer studies have explored children's semantic learning during independent reading (e.g., Joseph & Nation, 2018). A few have shown that semantic learning is associated with each of orthographic representations, word reading, and reading comprehension at the same point in time (Deacon, Mimeau, et al., 2019; Mimeau et al., 2018; Ricketts et al., 2011). Mimeau and her colleagues (Deacon, Mimeau, et al., 2019; Mimeau et al., 2018) also found that the relation between semantic learning and reading comprehension remained significant after controlling for several variables, including vocabulary, phonological awareness and word reading, pointing to a unique contribution of semantic learning to reading comprehension. In contrast, semantic learning was not further associated with word reading after taking control variables into account. Besides, no longitudinal studies investigating the relation between semantic learning and reading were identified in the literature, leaving the developmental nature of this relation untested.

The Present Study

In sum, on the basis of key theories of reading (e.g., Share, 1995; Perfetti & Hart, 2002), it can be predicted that orthographic and semantic learning play an important role in the development of orthographic representations, word reading, and reading comprehension. There is a small body of research that provides some evidence that orthographic learning is associated with orthographic representations and word reading longitudinally. In the present study, we extended this work critically by testing whether both of orthographic and semantic learning are related longitudinally to reading outcomes, embracing the potential and likely impact of semantics on reading (e.g., Quinn et al., 2015). Further, prior cross-sectional studies show specific links between orthographic learning and word reading and between semantic learning and reading comprehension. We added to this body of research by focusing not solely on word reading and reading comprehension, but also on orthographic representations, which are, together with phonological and semantic representations, at the very foundation of the reading process (Perfetti & Hart, 2002).

Following a group of children from Grade 3 to Grade 4, we hypothesised that each of children's orthographic and semantic learning in Grade 3 would be associated with their orthographic representations, word reading, and reading comprehension in Grade 4. Considering the central roles of orthography in word reading (e.g., Share, 1995) and of semantics in reading comprehension (e.g., Quinn et al., 2015), we expected the relations between orthographic learning and word reading and between semantic learning and reading comprehension to be particularly robust, remaining after the controls of non-verbal reasoning, working memory, vocabulary, and phonological awareness (and also word reading in the case of reading comprehension; see also Deacon, Mimeau, et al., 2019; Mimeau et al., 2018).

We further tested the directionality of these relations using autoregressors, that is, controlling for reading outcomes at Grade 3 (e.g., as in Deacon, Pasquarella, et al., 2019). Given the conservative nature of autoregressors and the well-known stability of reading skill over time, we expected several relations to become non-significant (Hulslander et al., 2010). Yet, since orthographic learning is deemed to lead directly to the mental representation of words' spelling (Share, 1995), we hypothesised that the relation between orthographic learning and orthographic representations would be the most likely to survive.

Method

The present study was part of a larger study on reading development. Mimeau et al. (2018) report on children's orthographic and semantic learning and reading outcomes at Grade 3. APA ethical standards were followed in the conduct of the study, and approval from

Dalhousie University's Social Sciences and Humanities Research Ethics Board was obtained.

Participants

We recruited children from six urban and five rural public schools in Nova Scotia, Canada, a largely Caucasian area (6% of the population is Aboriginal and another 6% is non-Caucasian; Statistics Canada, 2019). All participants had normal or corrected-to-normal vision. We focused on Grade 3 children because from that point, most children have sufficient decoding skills to understand texts (Chall, 1983; Hjetland et al., 2019), enabling them to gain new knowledge such as the spelling and meaning of new words. We obtained parental consent for 139 children, and we were able to follow up with 124 of these children when they were in Grade 4, on average 11.71 months later (SD = 0.43). The other 15 participants had moved (n = 14) or decided to withdraw from the study (n = 1). We conducted our analyses with the data for the participants assessed at both grades. Given the linguistic nature of the study, we further excluded four participants who did not speak English as their first language. This resulted in a sample of 120 children. The mean age of this sample was 8.17 years at Grade 3 (SD = 0.29) and 9.17 years at Grade 4 (SD = 0.29). There were 61 boys and 59 girls. Based on the locations of the schools, the mean household income was $77,740 (*SD* = $13,145), which is similar to the provincial average (Statistics Canada, 2019). As indicated by the standardised scores presented in Table 1, participants performed within the typical range on measures of word reading fluency and accuracy, non-verbal reasoning, working memory, and phonological awareness; their performance was slightly lower in reading comprehension, but still within the normal limits.

Materials

The reliabilities of all individual measures are presented in Table 1.

Learning Task

We measured orthographic and semantic learning within the same paradigm so that our results across these two aspects of learning would be comparable. We used a typical learning task, consisting of an exposure phase followed by post-tests (e.g., Ricketts et al., 2011). There were three orthographic learning post-tests and five semantic learning post-tests. In the present study, for semantic learning, we report only on the two post-tests that loaded properly on the Semantic Learning factor in Mimeau et al.'s (2018) measurement model (the loadings for these post-tests were .81 and .82, whereas the loadings for the excluded post-tests, definition and immediate and delayed matching, ranged from .27 to .37). All materials used in the learning task are available at https://osf.io/z4up8/.

Exposure Phase. ***Stories*.** Participants read out loud 12 stories, each about a new invention (e.g., a fish tank cleaner). The stories were based on those created by Wang et al. (2011). Each story contained five sentences and between 40 and 51 words, including four repetitions of one non-word that represented the invention. One of the sentences described the function of the invention explicitly (e.g. "The veap is used to clean fish tanks"). The stories were presented in sets of three in the same pre-randomised order for all participants. As participants read the stories, the experimenter provided feedback every time a word or non-word was mispronounced, skipped, or added. This feedback was intended to allow participants to fully understand the stories and know the correct pronunciation of the non-words, which was particularly important since the experimenter referred to them in the orthographic and semantic learning post-tests. Participants were not required to repeat the words or non-words for which they received feedback.

***Non-Words*.** Each non-word was four letters long and both started and ended with a consonant sound. All 12 non-words began with a different letter. The non-words had regular spellings, as per Rastle and Coltheart's (1999) rules. None of the non-words were listed in the Children's Printed Word Database (http://www.essex.ac.uk/psychology/cpwd), confirming their status as non-words for participants. Finally, each non-word contained a target sound that could be spelled in at least two ways (e.g., /i/, which can be spelled "ea" or "ee"). Each target sound was presented in two non-words, with a different spelling in each non-word (e.g., "veap" and "seef"). To control for any spelling preference, half of participants were given an alternative spelling of the non-words (e.g., "veep" and "seaf"). Supplementary

Table 1 presents the number of words sharing orthographical and phonological characteristics with each of the non-words.

Orthographic Learning Post-Tests. After reading each set of three stories, participants completed a spelling post-test. The experimenter read each of the three non-words from the set and asked participants to spell them on a sheet of paper (e.g., "Spell 'veap'"). Answers had to be identical to the non-words presented in the stories to be considered as correct. Within each set, the non-words were presented in the same pre-randomised order for all participants. Although administering the spelling post-test after each set of three stories might have slightly altered children's learning strategies for the following sets, we did so based on piloting and prior studies (e.g., Deacon, Pasquarella, et al., 2019) to maximise children's performance.

After reading the 12 stories, participants completed two orthographic choice post-tests: an immediate one right after the exposure phase to measure immediate recall and a delayed one a few days later to measure delayed retention (see the Procedure for more details). The two orthographic choice post-tests were identical. For each non-word, the experimenter showed four spellings to participants and asked them to choose the correct one (e.g., "Show me the spelling of 'veap'"). The non-words and the choices were presented in the same pre-randomised order for all participants. The distractors were the alternative spelling of the non-word (e.g., "veep") and two non-words that varied by one letter (e.g., "feap" and "feep"). Because the orthographic choice post-tests showed the written form of the non-words, they were administered after the spelling post-test, to limit any contamination between the post-tests.

No feedback was provided to participants on their answers during any of the orthographic learning post-tests. To reduce the number of analyses, we converted the raw scores on the three post-tests (spelling, immediate orthographic choice, and delayed orthographic choice) to *z* scores and combined them to create a composite score of orthographic learning. This method also improved the reliability of our measure of orthographic learning (Cronbach's α = .83 for the composite score).

Semantic Learning Post-Tests. After reading the 12 stories, participants completed two semantic choice post-tests: an immediate one and a delayed one, as for the orthographic choice post-tests. The two semantic choice post-tests were identical. For each non-word, the experimenter showed four pictures to participants and asked them to choose the correct one (e.g., "Show me the picture of a veap"). Eight pictures were taken or adapted from Wang et al.'s (2011) work, and the others were created using Adobe Illustrator. The non-words and the choices were presented in the same pre-randomised order for all participants. The distractors were an invention that used the same object as the non-word (e.g., a fish tank painter) and two inventions that used another object (e.g., a sock matcher and a sock fixer).

No feedback was provided to participants on their answers during any of the semantic learning post-tests. To reduce the number of analyses, we converted the raw scores on the two post-tests (immediate semantic choice and delayed semantic choice) to *z* scores and combined them to create a composite score of semantic learning. As for orthographic learning, this method improved the reliability of our measure of semantic learning (Cronbach's α = .73 for the composite score).

Reading Outcomes

Orthographic Representations. As an index of orthographic representations, we measured spelling ability with two tasks based on Olson et al.'s (1985) and Stanovich and West's (1989) work. In each task, participants were presented with 25 written sentences in which the last word was spelled in two different ways. The experimenter read the sentences out loud, and participants had to circle the correct spelling. The correct answer was presented first for half of the sentences. In the first task, the distractor was a non-word homophone (e.g., "At night, we go to sleap/sleep"). In the second task, the distractor was a word homophone (e.g., "I felt so full after all the dessert I ate/eight"). As such, the second task required the use of semantic information to choose the correct spelling, whereas the first one did not. The materials used in the two tasks are available at https://osf.io/z4up8/. To reduce the number of analyses, we converted the raw scores on the two tasks to *z* scores and combined them to create a composite score of orthographic representations at each grade (at Grade 3, Cronbach's α = .87 for the composite score; at Grade 4, Cronbach's α = .85 for the composite score).

Word Reading. We measured word reading with two standardised tasks. First, we measured word reading fluency with the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (Torgesen et al., 1999). In this task, participants were asked to read as many words as possible from a list of 104 words ordered in increasing difficulty. As per the manual, the task was discontinued after 45 seconds. Second, we measured word reading accuracy with the Word Identification subtest of the revised version of the Woodcock Reading Mastery Tests (Woodcock, 1998). In this task, participants were asked to read as many words as possible from a list of 106 words ordered in increasing difficulty. As per the manual, the task was discontinued after six consecutives errors. To reduce the number of analyses, we converted the raw scores on the two tasks to *z* scores and combined them to create a composite score of word reading at each grade.

Reading Comprehension. To measure reading comprehension, we used Level 3 of the Comprehension subtest of the fourth edition of the Gates-MacGinitie Reading Tests (MacGinitie et al., 2007). In this task, participants were given 35 minutes to read short texts silently and answer 48 multiple-choice questions assessing comprehension.

Control Variables

We included four control variables in our study: non-verbal reasoning, working memory, vocabulary, and phonological awareness. Each one has been found to be associated both with learning variables, such as orthographic and semantic learning, and reading variables, such as orthographic representations, word reading, and reading comprehension (e.g., Deacon, Mimeau, et al., 2019; Deacon, Pasquarella, et al., 2019; Mimeau et al., 2018; Ricketts et al., 2011). Therefore, it is important to control for these potential confounding variables in order to determine whether our learning variables are directly associated with our reading variables.

Non-verbal reasoning was measured with the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999), in which participants had to choose the missing pieces completing 32 pictures. Working memory was measured with the Digit Span subtest of the fourth edition of the Wechsler Intelligence Scale for Children (Wechsler, 2003), in which participants had to repeat 32 series of digits, half in the same order and half backwards. Vocabulary was measured with a shortened version of the third edition of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), in which participants had to choose the pictures that best illustrated 51 words (see Sparks & Deacon, 2015, for a validation of this shortened version in Grade 1 to 3 children). Phonological awareness was measured with the Elision subtest of the second edition of the Comprehensive Test of Phonological Processing (Wagner et al., 2013), in which participants had to repeat 34 words after deleting a syllable or phoneme.

Versions of the Tasks

For all tasks administered at both grades (i.e., orthographic representations, word reading, and reading comprehension), different versions were used at each grade. Since word reading and reading comprehension were assessed with standardised tasks, we used the available alternative forms of the tasks. For orthographic representations, we created two versions of each task that we compared in a pilot study. The pilot study included 37 English-speaking children enrolled in Grades 2 to 4 in a private school. The children completed each version of the tasks in one of two individual sessions. The order of the versions was counterbalanced across sessions. Based on the children's performance on each item of the tasks, we switched some items from one version to the other to make both versions of each task as equivalent in difficulty as possible. The corrected versions of each task, which we used in the present study, generated comparable performance scores: $t(36) = -0.66$, $p = .51$, for the task without semantics, and $t(36) = 0.00$, $p > .999$, for the task with semantics.

Procedure

Participants were tested in their school. At each of Grade 3 and Grade 4, they took part in two individual sessions and one group session. The second individual session and the group session were completed on average 3.01 days (*SD* = 2.69) and 5.26 days (*SD* = 2.91), respectively, after the first individual session. At both grades, the individual sessions included tasks on which we do not report in the present study.

In Grade 3, the first individual session consisted of, in order, the word reading fluency task, the exposure phase and immediate post-tests of the learning task (spelling after each set of three stories, then, in order, orthographic choice and semantic choice), and the non-verbal reasoning task. The second individual session consisted of, in order, the delayed post-tests of the learning task (in order, orthographic choice and semantic choice), the word reading accuracy task, the vocabulary task, the working memory task, and the phonological awareness task. The group session consisted of, in order, the orthographic representations task with no semantic component, the reading comprehension task, and the orthographic representations task with a semantic component.

In Grade 4, the first individual session consisted of the word reading fluency task and the second individual session consisted of the word reading accuracy task. The group session was the same as in Grade 3.

Results

We used IBM SPSS Statistics 23 to run our analyses, unless otherwise specified. We used the composite scores described in the Method section and the raw scores for the other variables. Data for one participant were removed from the analyses because the participant was unable to complete the learning task. Other missing data (< 1%) were missing completely at random according to Little's test, $\chi^2(54) = 48.86$, $p = .67$. Following Tabachnick and Fidell's (2007) recommendation to "replace a missing value with a value from a well-educated guess" (p. 66), these missing data were replaced by the participants' own data when other items or another similar task were available. When this was not possible, we used the mean from the remaining participants instead. We also examined our data for outliers. We identified four univariate outliers. We replaced these values by the next lowest or highest value instead of deleting participants to optimise our sample size. We identified no multivariate outlier. The residuals of all analyses were distributed normally as per the examination of Q-Q plots.

To ensure we categorised our key variables effectively, we conducted a confirmatory factor analysis in Mplus 7. We included an Orthographic Learning factor (three post-tests: spelling, immediate orthographic choice, delayed orthographic choice at Grade 3), a Semantic Learning factor (two post-tests: immediate semantic choice, delayed semantic choice at Grade 3), an Orthographic Representations factor (two tasks: without semantics, with semantics at Grade 4), a Word Reading factor (two tasks: fluency, accuracy at Grade 4), and reading comprehension as a single measure (at Grade 4). The model fitted our data well, $\chi^2(26) = 50.42$, $p = .003$, $\chi^2/\text{df} = 1.94$ (3 or lower indicates good fit), CFI = .97 (.95 or higher indicates good fit), SRMR = .05 (.08 or lower indicates good fit; Hu & Bentler, 1999; Iacobucci, 2010).

Descriptive statistics are presented in Table 1. In the learning task, children read 96% of the words and 78% of the non-words from the stories correctly. Furthermore, performance was above chance level in all orthographic and semantic choice post-tests (*p*s < .001; see Supplementary Table 2 for the proportion of correct answers in reading and the post-tests for each of the non-words separately). Supplementary Table 3 presents the correlations between the key variables. All correlations between the orthographic learning post-tests and the one between the semantic learning post-tests were significant. Within the reading measures (orthographic representations, word reading, and reading comprehension), all correlations were strong and significant, including those across grades. The correlations between the composite scores of orthographic and semantic learning at Grade 3 and the composite scores of reading outcomes at Grade 4 were also all significant (*r*s = .32–.63, *p*s < .001).

To determine whether orthographic and semantic learning at Grade 3 predicted orthographic representations, word reading, and reading comprehension at Grade 4, we conducted three hierarchical regression analyses. The dependent variable was either orthographic representations, word reading, or reading comprehension at Grade 4. For all regressions, we entered age, non-verbal reasoning, working memory, vocabulary, and phonological awareness at Grade 3 at Step 1 as control variables. For the regression involving reading comprehension, we also added word reading as a control at Step 1 to assess comprehension specifically, as in previous studies (e.g., Deacon, Mimeau, et al., 2019; Mimeau et al., 2018). Then, we entered orthographic learning and semantic learning at Grade 3 at Step 2. The results are summarised in Table 2. Our key findings are that children's orthographic learning at Grade 3 predicted unique variance in their orthographic representations and word reading, but not in their reading comprehension at Grade 4. Furthermore, children's semantic learning at Grade 3 predicted unique variance in their orthographic representations and reading comprehension, but not in their word reading at Grade 4.

We then performed a more stringent test of directionality: We conducted our three hierarchical regression analyses again including autoregressors, acknowledging that these are extremely conservative (see Hulslander et al., 2010). Specifically, in each regression analysis, we included the corresponding reading outcome at Grade 3 as an added control. The results are presented in Supplementary Table 4. Autoregressors accounted for between 46% and 87% of variance. With this added control, children's orthographic learning at Grade 3 still predicted unique variance in their orthographic representations but not in their word reading (or reading comprehension) at Grade 4. Similarly, children's semantic learning at Grade 3 still predicted unique variance in their orthographic representations but not in their reading comprehension (or word reading) at Grade 4.

Discussion

Prominent theories of reading imply that individual differences in children's capacity to learn words can support the development of different aspects of reading (Share, 1995; see also Ehri, 2020; Perfetti & Hart, 2002; Seidenberg & McClelland, 1989). In accordance with the developmental nature of these theories, the objective of the present study was to investigate, longitudinally, the relations between orthographic and semantic learning at

Grade 3 and orthographic representations, word reading, and reading comprehension at Grade 4. After controlling for age, non-verbal reasoning, working memory, vocabulary, and phonological awareness (and word reading when predicting reading comprehension), we found that orthographic learning at Grade 3 predicted orthographic representations and word reading at Grade 4; in parallel, semantic learning at Grade 3 predicted orthographic representations and reading comprehension at Grade 4. Our most stringent autoregressive analyses indicated that each of orthographic and semantic learning at Grade 3 predicted gains in orthographic representations from Grade 3 to Grade 4.

Our findings that earlier orthographic learning predicts later orthographic representations and word reading align with the few longitudinal studies investigating orthographic learning and reading outcomes (e.g., Deacon, Pasquarella, et al., 2019). These findings strengthen the evidence in support of Share's (1995) self-teaching hypothesis. In particular, our longitudinal design provides direct support for the developmental implications of the hypothesis.

In addition, our findings that earlier semantic learning is associated with later orthographic representations and reading comprehension extend prior research with concurrent designs. Indeed, we began exploring a key and missing piece of evidence by investigating the *longitudinal* relations between semantic learning and reading. Our findings suggest that the self-teaching hypothesis (Share, 1995) could be extended to semantic learning (see also Mimeau et al., 2018). In other words, we think that individual differences in children's capacity to learn both the spelling and the meaning of words could support different aspects of their reading development.

The results of our autoregressive analyses highlight the contribution of each of orthographic and semantic learning to the development of orthographic representations of words over time. These results align with Deacon, Pasquarella, et al.'s (2019) finding that orthographic learning at Grades 2 and 3 predicts one-year gains in orthographic representations and word reading, taken together. This emerging body of research points in the direction of a potential causal relation from orthographic and semantic learning to orthographic representations (see Selig & Little, 2012). For orthographic learning, this idea is in line with Share's (1995) argument that orthographic learning leads to efficient word reading through the formation of a mental store of orthographic representations; this store is what our measures of orthographic representations were tapping. The robustness of the contribution of semantic learning to gains in orthographic representations, however, is more novel and unexpected. One possibility suggested by some theories of reading is that the connections between the different aspects of word representations—phonological, orthographic, and semantic—are particularly strong (e.g., Seidenberg & McClelland, 1989). Learning the meaning of new words during independent reading would then enable the mental formation of semantic representations of these words but also reinforce their orthographic representations.

As we consider these explanations, we bear in mind the highly conservative nature of the autoregressive analyses we conducted (see Hulslander et al., 2010); as such, the null effects we observed do not necessarily mean that relations do not exist. In particular, the standardised tests we used to measure word reading and reading comprehension were very stable in time (*r*s = .71–.90) and thus left little variance to be explained by other factors. Further, as argued by Hulslander et al. (2010), it might be that beyond the first years of learning to read, measures of word reading and reading comprehension already include the effect of causal factors. For example, word reading at Grade 3 could include pre-existing effects of orthographic learning. Controlling for word reading at Grade 3 would thus control for orthographic learning too, making it unlikely that our measure of orthographic learning contributes further to word reading at Grade 4.

Beyond the contribution of our study to theories such as the self-teaching hypothesis (Share, 1995), our findings also have important educational implications. Indeed, they point to the potential value of reading programs that foster the development of cognitive and behavioral strategies to support word learning during independent reading. For instance, children could be encouraged to focus on the spelling of new words (e.g., White, 2005) and on their meaning (e.g., Cantrell et al., 2010) as they read independently. Of course, the effectiveness of such programs would need to be tested, but they would respond directly to encouragement to shift instruction towards learning (Kilpatrick, 2018).

The conclusions we draw from our results should be considered alongside our study's limitations. First, we provided feedback to children as they were reading the stories in the learning task to limit the influence of decoding on results. That said, this decision reduces the comparability of the learning task to children's natural independent reading and to the original self-teaching task (Share, 1995). Second, although we report some data on the orthotactic probabilities of our non-words in Supplementary Table 1, we did not control for these characteristics in the present study. This should be borne in mind when interpreting our results, given that orthotactic (and phonotactic) probabilities have been found to affect orthographic learning (Apel et al., 2006). Third, while we included a measure of nonverbal intelligence in our analyses, we did not control for general learning ability more specifically, which could explain the associations we found with orthographic and semantic learning. Further studies are needed to disentangle the role of word learning capacity versus general learning ability in reading development. Future research could also measure orthographic representations with a standardised spelling test, instead of experimental tasks as we did, to make all three reading outcomes more comparable.

Fourth, although the longitudinal design of our study helped identify potential causal factors involved in the development of reading, only experimental designs truly test causality. It would also be interesting in the future to investigate the opposite direction of the relations, that is, whether earlier reading skill can predict later word learning capacity. Fifth, our sample size was appropriate for regression analysis and indeed enabled us to identify significant longitudinal associations, even after including stringer autoregressors. It was not, however, sufficient to perform structural equation modeling that might untangle the complex relations between orthographic and semantic learning and different aspects of reading proposed in theories of reading development. Using structural equation modeling would also have the advantage of reducing measurement error, which could not be done with composite scores. Finally, the fact that we conducted several regression analyses increased Type I error, so future studies are needed to replicate and confirm our findings.

In conclusion, we found that individual differences in children's capacity to learn the spelling and meaning of new words during independent reading can predict how well they will read words and understand texts, respectively, one year later; these individual differences in children's word learning capacity can also predict how much their mental representations of words' spelling will improve over one year. These findings bring support to and extend prominent theories on the role of orthographic learning and semantics in reading, such as the self-teaching hypothesis (Share, 1995) and the lexical quality hypothesis (Perfetti & Hart, 2002; see also Ehri, 2020; Seidenberg & McClelland, 1989). Notably, the longitudinal nature of our study provides a first test of the developmental implications of these theories and their potential applicability to classroom instruction, encouraging the use of learning strategies to support word learning during independent reading.

References

Apel, K., Wolter, J. A., & Masterson, J. J. (2006). Effects of phonotactic and orthotactic probabilities during fast mapping on 5-year-olds' learning to spell. *Developmental Neuropsychology*, *29*(1), 21–42. https://doi.org/10.1207/s15326942dn2901_3

Cantrell, S. C., Almasi, J. F., Carter, J. C., Rintamaa, M., & Madden, A. (2010). The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers. *Journal of Educational Psychology*, *102*(2), 257–280. https://doi.org/10.1037/a0018212

Chall, J. S. (1983). *Stages of reading development*. McGraw-Hill.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *33A*(4), 497–505. https://doi.org/10.1080/14640748108400805

Deacon, S. H., Mimeau, C., Chung, S. C., & Chen, X. (2019). Young readers' skill in learning spellings and meanings of words during independent reading. *Journal of Experimental Child Psychology*, *181*, 56–74. https://doi.org/10.1016/j.jecp.2018.12.007

Deacon, S. H., Pasquarella, A., Marinus, E., Tims, T., & Castles, A. (2019). Orthographic processing and children's word reading. *Applied Psycholinguistics*, *40*(2), 509–534. https://doi.org/10.1017/S0142716418000681

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the Peabody Picture Vocabulary Test, third edition (PPVT-III): Form IIIA and Form IIIB*. American Guidance Service.

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, *55*(S1), S45–S60. https://doi.org/10.1002/rrq.334

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/10.1037/dev0000512

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Hjetland, H. N., Lervåg, A., Lyster, S.-A. H., Hagtvet, B. E., Hulme, C., & Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology*, *111*(5), 751–763. https://doi.org/10.1037/edu0000321

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulslander, J., Olson, R. K., Willcutt, E. G., & Wadsworth, S. J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading*, *14*(2), 111–136. https://doi.org/10.1080/10888431003604058

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, *166*, 190–211. https://doi.org/10.1016/j.jecp.2017.08.010

Kilpatrick, D. A. (2018). Incorporating recent advances in understanding word-reading skills into specific learning disability diagnoses: The case of orthographic mapping. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 947–972). The Guilford Press.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2007). *Gates-MacGinitie Reading Tests, fourth edition: Level 3, Forms S&T: Directions for administration*. Riverside Publishing.

Mimeau, C., Ricketts, J., & Deacon, S. H. (2018). The role of orthographic and semantic learning in word reading and reading comprehension. *Scientific Studies of Reading*, *22*(5), 384–400. https://doi.org/10.1080/10888438.2018.1464575

Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 4, pp. 1–64). Academic Press.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Studies in written language and literacy: Vol. 11. Precursors of functional literacy* (pp. 189–213). John Benjamins.

Pritchard, S. C., 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*, *42*(3), 722–770. https://doi.org/10.1111/cogs.12571

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, *86*(1), 159–175. https://doi.org/10.1111/cdev.12292

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 482–503. https://doi.org/10.1037/0096-1523.25.2.482

Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47–70. https://doi.org/10.1080/10888438.2011.536129

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. http://dx.doi.org/10.1037/0033-295X.96.4.523

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (p. 265–278). The Guilford Press.

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

Sparks, E., & Deacon, S. H. (2015). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics*, *36*(2), 299–321. https://doi.org/10.1017/S0142716413000246

Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education*, *189*(1–2), 23–55. https://doi.org/10.1177/0022057409189001-204

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402–433. https://doi.org/10.2307/747605

Statistics Canada (2019, June 18). *Census profile, 2016 census*. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How word meaning influences word reading. *Current Directions in Psychological Science*, *24*(4), 322–328. https://doi.org/10.1177/0963721415574980

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Examiner's manual: Test of Word Reading Efficiency (TOWRE)*. PRO-ED.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing, second edition (CTOPP-2): Examiner's manual*. PRO-ED.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, *109*(1), 39–57. https://doi.org/10.1016/j.jecp.2010.11.005

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI): Manual*. Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children, fourth edition (WISC-IV): Administration and scoring manual*. Psychological Corporation.

White, T. G. (2005). Effects of systematic and strategic analogy-based phonics on grade 2 students' word reading and reading comprehension. *Reading Research Quarterly*, *40*(2), 234–255. https://doi.org/10.1598/RRQ.40.2.5

Wolter, J. A., Self, T., & Apel, K. (2011). Initial mental graphemic representation acquisition and later literacy achievement in children with language impairment: A longitudinal study. *Journal of Learning Disabilities*, *44*(6), 543–555. https://doi.org/10.1177/0022219410392042

Woodcock, R. W. (1998). *Woodcock Reading Mastery Tests–Revised: Forms G and H: Examiner's manual*. American Guidance Service.

Table 1

Descriptive Statistics for Orthographic and Semantic Learning, Orthographic Representations, Word Reading, Reading Comprehension, and Control Variables

Variable	Max	Grade 3			Grade 4		
		M	*SD*	Reliability	*M*	*SD*	Reliability
Orthographic learning							
Spelling	12	5.91	3.11	.76	—	—	—
Immediate orthographic choice	12	8.87	2.22	.58	—	—	—
Delayed orthographic choice	12	8.54	1.86	.33	—	—	—
Semantic learning							
Immediate semantic choice	12	9.34	1.94	.53	—	—	—
Delayed semantic choice	12	9.76	1.73	.54	—	—	—
Orthographic representations							
Without semantics	25	19.27	3.60	.75	20.93	3.04	.72
With semantics	25	19.92	3.83	.79	22.54	2.85	.78
Word reading							
Fluency	104	57.85	13.18	.95 [a]	63.01	12.07	.93 [a]
Standard score (*M* = 100, *SD* = 15)		103.74	13.81		102.24	12.51	
Accuracy	106	60.86	11.33	.97 [a]	67.61	12.31	.97 [a]
Standard score (*M* = 100, *SD* = 15)		106.36	10.65		104.25	11.28	
Reading comprehension	48	26.78	11.26	.93 [a]	31.53	9.84	.93 [a]
Normal curve equivalent (*M* = 50, *SD* = 21.06)		42.98	21.94		40.26	20.63	
Control variables							
Non-verbal reasoning	32	15.29	6.45	.93 [a]	—	—	—
T score (*M* = 50, *SD* = 10)		47.55	10.80		—	—	
Working memory	32	13.17	2.31	.86 [a]	—	—	—
Scaled score (*M* = 10, *SD* = 3)		9.18	2.48		—	—	
Vocabulary	51	32.04	4.98	.78	—	—	—
Phonological awareness	34	23.43	5.78	.91 [a]	—	—	—
Scaled score (*M* = 10, *SD* = 3)		9.22	2.54		—	—	

Note. Max = maximum score.

[a] These reliabilities come from the manual; the other ones were calculated from our data (Cronbach's alphas).

Table 2

Results of the Hierarchical Linear Regressions Predicting Reading Outcomes at Grade 4 From Orthographic Learning and Semantic Learning at Grade 3

Predictor [a]	Orthographic representations			Word reading			Reading comprehension		
	B (*SE*)	β	ΔR^2	*B* (*SE*)	β	ΔR^2	*B* (*SE*)	β	ΔR^2
Step 1			.29***			.41***			.60***
Age	0.20 (0.22)	.06		−0.01 (0.22)	.00		−0.75 (2.09)	−.02	
Non-verbal reasoning	−0.01 (0.01)	−.06		0.01 (0.01)	.04		0.07 (0.10)	.05	
Working memory	0.08 (0.03)	.21**		0.07 (0.03)	.17*		0.52 (0.29)	.12	
Vocabulary	0.00 (0.01)	.01		0.04 (0.01)	.20*		0.39 (0.14)	.20**	
Phonological awareness	0.01 (0.01)	.09		0.04 (0.01)	.25***		−0.03 (0.13)	−.02	
Word reading							6.18 (0.91)	.58***	
Step 2			.21***			.10***			.02
Orthographic learning	0.49 (0.09)	.47***		0.42 (0.09)	.39***		−0.42 (0.91)	−.04	
Semantic learning	0.20 (0.08)	.20*		−0.03 (0.08)	−.03		1.76 (0.75)	.17*	

Note. The reported statistics are from the final models that included all variables.

[a] All predictors were measured at Grade 3, except for word reading, which was measured at Grade 4.

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