

The Word Nerds project: Findings from a research–practice partnership focused on spelling instruction

Laura S. Tortorelli 

Department of Teacher Education, Michigan State University, East Lansing, Michigan, USA

Lori Bruner 

Department of Teacher Education, Michigan State University, East Lansing, Michigan, USA

Learning to spell in English requires the integration of general and specific word knowledge. This paper describes the ‘Word Nerds’ project, a research–practice partnership consisting of two researchers from a large public university and 17 elementary teachers in seven school districts in the United States. The collaboration was formed to study variation in instructional practice among teachers using the *Words Their Way* programme and address teacher-generated questions related to how children learn to spell words. This paper describes how stakeholders worked together to understand (1) what grouping and organisational structures teachers use for spelling instruction, (2) the extent to which elementary students use analogy to spell unknown words that share spelling patterns with known words and (3) the affordances and challenges of a partnership approach to educational research. Data collection took place over one school year. Data sources included teacher focus groups, classroom observations and student formative spelling assessment data ($n = 178$). Analyses included descriptive statistics, t -tests and multilevel modelling, nesting spelling items in students in classrooms. Findings indicated that classroom observation scores for grouping, teacher talk, student-to-student talk, reflection and student engagement during spelling lessons varied depending on format (whole group, small group or pair/individual). Students showed evidence of spelling by analogy, but this strategy was associated in part with the semantic difficulty of the target words. Instructional practices for spelling vary based on teachers’ priorities, and teachers need support to implement and prioritise evidence-based practices. Conducting research in partnership allowed the researchers to observe and understand variation in implementation of a popular spelling programme and to develop and answer research questions that mattered to teachers; however, it also limited the size, scope and location of the project.

Keywords: orthography, research–practice partnerships, spelling

Highlights

What is already known about this topic

- Children can read and spell words by analogising familiar spelling patterns from known to unknown words.
- Evidence-based spelling instruction focuses on common orthographic and morphological patterns that help students use analogy to spell and form high-quality lexical representations in memory.
- Implementation of evidence-based spelling instruction is challenging for teachers, and there is little research examining how popular commercial spelling programmes are used in classrooms.

What this paper adds

- Classrooms engaged in small-group spelling instruction included more teacher talk; classrooms engaged in individual instruction were highly differentiated; and classrooms engaged in whole-group instruction included more student-to-student talk and engagement.
- Student data showed evidence of spelling by analogy, but students were less likely to spell more semantically challenging words correctly and made more errors in spelling those words, even when orthographic difficulty was held constant.
- The research–practice partnership allowed researchers to document and explain variability in spelling instruction and address teachers’ questions about spelling development.

Implications for theory, policy or practice

- Teachers adapt spelling instruction to address their priorities, and different organisational models prioritise direct instruction, differentiation, standards-based content, or peer collaboration.
- The ability to analogise spelling patterns from known to unknown words may depend on linking orthographic knowledge to vocabulary knowledge.
- Teachers need support and community to implement and iterate best practices for spelling instruction.

INTRODUCTION

Spelling ability plays an important role in literacy development; however, teachers often struggle to implement research-based spelling instruction (e.g., Fresch, 2003, 2007; Graham et al., 2008). This paper describes the ‘Word Nerds’ project, a *research–practice partnership* (Penuel & Gallagher, 2017) consisting of researchers from a large public university and literacy leaders and 17 elementary teachers in seven school districts in the

United States. The goals of the project were to document spelling instruction in elementary classrooms using *Words Their Way*, a popular spelling programme, and answer teacher-generated questions about how their students use analogy to generalise their spelling knowledge to untaught words.

Political context and impetus for the partnership

Literacy instruction in the United States takes place in a high-stakes accountability environment from the first day of kindergarten (e.g., Bassok et al., 2016). In 2015, Michigan fell to 41st out of 50 U.S. states on national assessments, and 50% of third graders were scoring below proficient on state exams, indicating that they had not mastered the grade-level literacy content standards (Michigan Department of Education, 2017; National Center for Education Statistics, 2016). In response, Michigan enacted legislation that enabled schools to retain these students (i.e., hold them back in third grade for a second year before moving them up to fourth grade). Given that this policy could potentially affect half of the third-grade students in the state, there was an urgent need for research conducted with school leaders and teachers to understand *what* was happening in elementary literacy instruction and *how* and *for whom* it needed to be improved. Literacy leaders identified *Words Their Way* as commonly used spelling programme being adopted in many districts to support early literacy. *Words Their Way* is among Pearson's best-selling educational books in the United States, but little research was available on its implementation and effectiveness. Thus, we found an area for collaborative work that could strengthen local practice and contribute more broadly to the field. The authors worked with elementary teachers using *Words their Way* over one school year and included monthly focus groups, classroom observations and the collection and analysis of student spelling data at the beginning and end of the year.

Theoretical framework

Spelling ability is an important aspect of overall literacy (e.g., reading, writing, listening and speaking) development (Graham & Santangelo, 2014; Treiman, 2017; Treiman et al., 2019). High-quality word representations in memory underpin both word recognition and spelling and provide a critical foundation for higher-level literacy skills (e.g., text comprehension and written composition; Perfetti & Hart, 2002). A high-quality representation is fully specified in orthographic memory and tightly bound to the word's phonological and semantic identities, allowing for rapid, reliable access during reading and writing tasks. Because spelling is an expressive task, 'variation in lexical quality can be observed in spelling' (Perfetti & Hart, 2002, p. 140).

For children learning to write in English, spelling depends on the ability to integrate general and specific word knowledge (Share, 1995; Treiman, 2017; Treiman & Kessler, 2014; Venezky, 1999). Children use phonemic awareness, alphabet knowledge, and phonics and spelling knowledge to phonologically recode specific words, embedding them in memory for future use (Share, 1995). As children develop larger lexicons of known words, the process of recoding becomes increasingly lexicalised, focusing on larger patterns and units that frequently appear within words (e.g., *make-take-lake*; *eat, eating, eater*; Ehri, 2005; Goswami, 1999; Share, 1995; Treiman, 2017). For example, the word *make* could plausibly be spelled *mace*, *maik* or *mayk*, but the a-consonant-e pattern is commonly used with a

final *k*. Morphemes, units of meaning within words, are typically spelled the same way across words, even when they represent different sounds (e.g., the final *t* in *eat* and *eater*; Levesque et al., 2021; Treiman, 2017; Treiman & Kessler, 2014). Thus, a child's vocabulary knowledge and the semantic difficulty of specific words play a crucial role in distinguishing between plausible spellings and storing and retrieving words in memory. In reading, this process has been demonstrated through studies of *set for variability*, which indicate that word recognition depends on a reader's ability to flexibly map ambiguous spellings to known words in their oral vocabularies (e.g., does *mow* rhyme with *snow* or *cow*?; Elbro et al., 2012; Tunmer & Chapman, 2012). Some features of words, including word frequency (how often a word appears in text) and length in letters, indicate words that are both more difficult to spell and less likely to be part of children's oral vocabularies (Hiebert et al., 2019; Share, 1995). Additional semantic features that influence how likely a child is to know the meaning of a word include age of acquisition (the age at which a word is typically known in oral language), concreteness (how easily a reader can generate a mental image for the word) and part of speech (Hiebert et al., 2019). Children with larger vocabularies show greater ability to read words with ambiguous vowel patterns (Ricketts et al., 2007), and learning to spell words supports both orthographic learning (Conrad et al., 2019; Shahar-Yames & Share, 2008) and vocabulary learning (Ricketts et al., 2009; Rosenthal & Ehri, 2008). Thus, learning to spell words supports the construction of high-quality lexical representations, which are an important driver of literacy acquisition (Conrad et al., 2019; Ouellette et al., 2017; Treiman et al., 2019). Weaker spellers not only spell fewer words correctly, but their spelling knowledge is less integrated with their phonological and vocabulary knowledge (Perfetti & Hart, 2002; Treiman, 2017), and as a result, they will be less able to use analogy from known words to spell similar unknown words.

Spelling instruction

Despite the importance of spelling, research indicates that teachers struggle to follow evidence-based practices for spelling instruction and that practices vary widely across classrooms (Fresch, 2007; Ganske & Jocius, 2013; Graham et al., 2008; Schlagal, 2002). As a result, examining spelling instruction is critical to supporting literacy development. Approaches to spelling instruction summarised by Treiman (2017) include phase-and stage-based approaches, in which instruction is differentiated to target more sophisticated features over time, and the integration of multiple patterns approach (Treiman & Kessler, 2014), in which children are taught to use both sound-spelling correspondences and common contextual and morphological patterns to spell words. The *Words Their Way* programme draws on both approaches, with its central tenets being (1) differentiated instruction based on developmental spelling stages (Henderson, 1990) and (2) systematic study of words in groups that share spelling patterns (Bear et al., 2019). An overview of the *Words Their Way* spelling stages and instruction are presented in Table 1. Students compare, contrast and sort words into categories using the phonological, orthographic, contextual and morphological information that is necessary for learning to spell in English (Treiman, 2017; Treiman & Kessler, 2014). *Words Their Way* became a focus of the project because (1) it is theoretically based, as described above, and (2) it is widely used in the United States and the partner districts.

Table 1. Overview of the *Words Their Way* development spelling stages.

Targeted features for instruction	Examples	Common core/Michigan State standards for spelling in Grades K-5
Emergent (PreK to early first grade)		
Letter sounds	Pictures of words that start with s versus z	
Initial consonants		
Letter–name alphabetic (kindergarten–Early Grade 2)		
Initial consonants	pill will hill	Kindergarten L2c. Write a letter or letters for most consonant and short-vowel sounds (phonemes).
Short vowels	bag beg big pat bend last	Kindergarten L2d. Spell simple words phonetically, drawing on knowledge of sound–letter relationships.
Consonant blends and digraphs	ship chip step trap	
Within word pattern (Late Grade 1 to Middle Grade 4)		
Long vowel patterns	woke broke stone phone boat float loan moan	First Grade L2d. Use conventional spelling for words with common spelling patterns and for frequently occurring irregular words.
Diphthongs and other ambiguous vowel patterns	cow town out house head lead lean meal	First Grade L2e. Spell untaught words phonetically, drawing on phonemic awareness and spelling conventions. Second Grade L2d. Generalise learned spelling patterns when writing words (e.g., cage → badge; boy → boil).
Syllables and affixes (Grades 3 to 6)		
Compound words	snowplow racecar outdoor	Third Grade L2e. Use conventional spelling for high-frequency and other studied words and for adding suffixes to base words (e.g., <i>sitting</i> , <i>smiled</i> , <i>cries</i> , <i>happiness</i>).
Open and closed syllables	pilot pillow robot goblet	Third Grade L2f. Use spelling patterns and generalisations (e.g., word families, position-based spellings, syllable patterns, ending rules and meaningful word parts) in writing words.
Base words and affixes	waited spelled played stopped dotted patted raced biked carried hurried	Fourth Grade L2d. Spell grade-appropriate words correctly, consulting references as needed. Fifth Grade L2e. Spell grade-appropriate words correctly, consulting references as needed.
Derivational relations (Grades 5 and up)		
Greek and Latin roots and stems	act action distinct distinction select selection predict prediction	

Despite wide use in classrooms since in 1996, *Words Their Way* has rarely been the subject of research. In 2013, a review by the U.S. Institute of Educational Sciences' What Works Clearinghouse found no efficacy studies of *Words Their Way*, and our own review found no peer-reviewed efficacy studies since 2013. One reason may be that *Words Their Way* is not a scripted programme, and spelling instruction varies widely. Thus, the Word Nerds project was conceptualised as a research–practice partnership to work collaboratively with teachers and literacy leaders to understand spelling instruction from a teacher-centred perspective and to document the variation in implementation of the *Words Their Way* programme (Bryk et al., 2015).

Research questions

The project began in 2017 with a period of 'rapid ethnography' (focus groups and informal observations) to inform the co-design process (Penuel & Gallagher, 2017). In initial focus groups, the group named the project 'Word Nerds' and research questions were co-designed. The focus groups revealed that there was significant variation in how teachers were organising and implementing spelling instruction with *Words Their Way*, as they attempted to integrate their materials, state standards and logistical concerns. They also expressed curiosity and frustration about student progress and whether or not their instruction transcended the taught words and improved students' ability to spell words with similar patterns, a central tenet of the programme. Out of these discussions, we developed two research questions to guide our work, one focusing on teaching and one focusing on student learning:

- 1 What grouping and organisational structures do teachers use for spelling instruction and what advantages and disadvantages does each structure offer?
- 2 To what extent do elementary students use analogy to spell unfamiliar words?

As researchers embarking on our first research–practice partnership, we had the additional aim of learning about the affordances and challenges of taking a partnership approach to educational research.

METHOD

This paper focuses on an early-stage research–practice partnership, in which researchers and educational agencies work together to 'take stock of current practices and contexts' (Penuel & Gallagher, 2017, p. 81). As such, the project was directed by a working group that co-designed the research project and worked together to collect and analyse data on spelling instruction (Coburn & Penuel, 2016).

Participants

Working group participants. The working group included 1 faculty member from a local research university, 2 research assistants, 17 teachers, 4 literacy coaches and 3 district literacy coordinators representing 13 schools in 7 school districts. Three teachers taught kindergarten (the first year of formal school in the United States), six taught first grade, three taught second grade, four taught third grade and one taught fourth grade. One

participant was male and one was Black; all others were white women. Fifteen teachers had master’s degrees and two had 4-year bachelor’s degrees. Teaching experience ranged from 4 to 25 years and experience teaching *Words Their Way* from 0 to 5 years. All participants had participated in a 2-day professional development on *Words Their Way* in spring 2017, which provided an overview of developmental spelling stages (see Table 1), the recommended scope and sequence, instructional routines and organisation, and practice for analysing student spelling data, grouping students for instruction and selecting learning goals and materials based on spelling stage. All teachers participated in the focus groups and observations that provided data to answer Research Question 1.

Student participants. To answer Research Question 2, spelling data were collected in fall 2017 and spring 2018 in 10 Grade 1–3 classrooms. This sample included 123 first-grade students (age 6–7 in the United States), 52 second-grade students (age 7–8) and 30 third-grade students (age 8–9). To minimise disruption and additional testing in participating classrooms, the project was designed to use data that teachers were already collecting from students. *Words Their Way* provides the Primary Spelling Inventory (PSI) as a spelling assessment for Grades 1–3; however, alternative assessments are provided for Grades 3 and up. Thus, kindergarten and fourth-grade teachers and some third-grade teachers did not participate in student data collection using the PSI. The student participants were 49% female, 80% white, 9% Latinx and 5% Black. Eight percent of students were receiving English as a Second Language service (provided for students whose first/home language is not English), 8% were receiving special education services (provided for students with identified disabilities) and 39% were receiving free and reduced priced lunch (provided to students living in poverty in the United States). To evaluate students’ baseline literacy performance, we analysed beginning of the year Dynamic Indicators of Basic Early Literacy Skills (DIBELS) scores. DIBELS is a widely used literacy assessment in the United States and includes subtests focused on letter recognition, phonemic segmentation, nonsense word reading and reading fluency and is used to identify students who are at risk of failing state-level literacy exams. At the beginning of project, 60% of first-grade participants, 85% of second-grade participants and 75% of third-grade participants met grade-level DIBELS benchmarks.

Table 2. Timeline of research activities.

Activity	Dates	Instruments
Focus groups	June 2017	Recordings
	August 2017	Contemporaneous notes
	October 2017	
	December 2017	
	February 2018	
	April 2018	
	June 2018	
Classroom observations	October 2017	Classroom Observation Tool
	February–March 2018	
	May–June 2018	
Student spelling data collection	September–October 2017	Primary Spelling Inventory
	May 2018	Alternative Primary Spelling Inventory

Data sources and procedures

The research activities are summarised in Table 2 and described below.

Focus groups. Focus groups involved the first author guiding a free-ranging discussion at bi-monthly intervals throughout the 2017–2018 school year (Flores & Alonso, 1995). Teachers discussed challenges in implementing *Words Their Way*, the first author presented the project data and analyses as it was available and the group discussed possible explanations, implications and directions for further research. All focus groups were recorded and contemporaneous notes were taken.

Classroom observations. The researchers travelled to the classrooms of participating teachers to observe spelling instruction in the fall, winter and spring. Classroom observation data were collected using the *Words Their Way Classroom Observation Tool* (COT; Gehsmann & Bear, 2013). The Classroom Observation Tool was developed to evaluate classrooms using the *Words Their Way* programme. The Classroom Observation Tool includes 10 categories of data: grouping, preparation and organisation of materials, teacher talk, facilitation of student-to-student talk, extension and transference, instructional routines, student reflection, notebook use, student engagement, and teacher knowledge and classroom management. Each category can be scored 0–3, with 3 representing highly proficient and 0 representing ‘not observed’. Inter-rater reliability is .84 and Cronbach’s alpha is .90 (Gehsmann et al., 2012). The researchers double-scored 25% of the observations independently. Following each joint observation, the scores were discussed and agreement was reached. The observations encompassed whatever the participating teacher described as the spelling lesson for the day and varied from 10 minutes to over an hour.

Student data. Participating teachers collected student consent forms, demographic data and spelling assessment data to share with the research team. At the child level in Grades 1–3, data sources included formative spelling assessment data from the PSI. The PSI is a group-administered spelling test that includes 26 items beginning with simple three letter words (*fan*) and ending with two-syllable words that require spelling changes to add affixes (*riding*). Item difficulty on the PSI ranged from *fan*, which 96.3% of a sample of 647 elementary school students can spell correctly to *clapping*, which 16.1% spell correctly (Sterbinsky, 2007). Item discrimination (an index of how differently high and low scoring students perform on a particular item) ranged from *fan* (6.3, meaning little difference between high and low performers) to *shine* (77.7, meaning large differences between high and low performers). Test–retest reliability was .90 and Cronbach’s alpha, a measure of internal consistency of the overall inventory, was .93, both of which are considered to be high/acceptable. PSI scores are significant predictors of performance on standardised tests of reading in Grades 1–3 (Sterbinsky, 2007). Students do not study for the PSI and are encouraged to spell the words as best as they can to reveal their knowledge about common orthographic patterns in English.

To test students’ ability to analogise, teachers also administered a researcher-created alternative form, called the Alternative Primary Spelling Inventory (APSI). At the request of the teachers in the working group, the first author created this assessment by matching each word on the PSI with another English word that shared the final rime and any suffixes (*dig* vs *jig*, *shade* vs *spade* and *shouted* vs *spouted*). Rimes were matched because rimes represent the most consistent orthographic–phonological units in English (Treiman et al., 1995),

and teaching students to recognise and use rime-based contextual information to spell words is a major focus of *Words Their Way* instruction. Items for the APSI were chosen to follow the developmental sequence of spelling features described in Henderson's (1990) research and taught in *Words Their Way*. According to this scope and sequence, initial consonants, consonant digraphs and consonant blends are all mastered in the initial stages of learning to spell and are target features in kindergarten and first grade. The larger rime and affix patterns that were held consistent between the PSI and APSI included vowel digraphs, diphthongs, r-controlled vowels and morphological affixes, which are learned later in development. The words used in the APSI were analysed for semantic difficulty (length, frequency, concreteness, age of acquisition and part of speech) using the Word Zone Profiler tool (Brysaert et al., 2014; Hiebert, 2012; Hiebert et al., 2019; Kuperman et al., 2012).

Teachers administered the PSI and APSI on different days within a 2-week period at the beginning and end of the school year. Teachers provided copies of the student answer sheets to the research team. The authors and a research assistant scored the samples, double-scoring 40% to establish inter-rater reliability. Each response on the PSI and APSI was scored correct (1) or incorrect (0), for a total score out of 26 possible points on each form. The APSI responses were also scored using Levenshtein distance (Levenshtein, 1966) to capture the difference between the provided spelling and an accurate spelling, assigning one point for each letter that would need to be added, subtracted or substituted to change the provided spelling into a correct spelling.

Data analysis

Teacher Classroom Observation Tool scores were analysed using descriptive statistics. Student PSI and APSI scores were analysed using descriptive statistics, paired sample *t*-tests and three-level multilevel regression using HLM-7 software, nesting item scores within students within classrooms (Raudenbush & Bryk, 2002; Raudenbush et al., 2011). At each stage of data collection and analysis, the data were presented to the focus groups to question, interpret and contextualise.

RESULTS

Focus group data

The research–practice partnership approach values and integrates teachers' and school leaders' professional knowledge of the local context (Erickson, 2014; Henrick et al., 2016). The group identified areas of alignment between school, district and state initiatives and their spelling practice, areas of tension and practical concerns about their spelling instruction.

Alignment. Teachers identified *Words Their Way* as a welcome form of support to enact the state mandates related to early literacy in their classrooms. They typically offered writing instruction through workshop models that focused on big-picture skills, for example, adding details to a narrative, but they understood that their students' lack of transcription skills hampered their ability to fully engage in these activities. Several teachers described having relied on older colleagues to support their phonics and spelling instruction in the

past, but noted that those colleagues had retired, and they felt their schools now lacked this knowledge at an institutional level. They appreciated that *Words Their Way* provided a framework for differentiation and a scope and sequence for spelling content, but felt they needed more support to implement and to build effective instructional routines.

Tensions. Teachers found that the principles embodied in *Words Their Way* were in tension with other top-down directives. In particular, they worried that by differentiating instruction, they were not moving all their students towards state standards that require mastery of particular spelling concepts in each grade level (see Table 1). Many teachers felt pressure to teach all students standards-aligned content and felt that teaching them ‘at their level’ was inequitable in the context of the recently passed ‘Third Grade Reading’ laws. Teachers also felt that grouping students for spelling instruction created logistical challenges. They struggled to keep everyone productively occupied while working with one group and keeping track of materials and lesson plans for multiple groups every week.

Teachers were also concerned that their spelling instruction was too divorced from the rest of the literacy curriculum. They believed that children were able to read and spell the words they were sorting but were not able to apply that knowledge in the rest of their reading and writing activities. They felt that the PSI captured incomplete data about their students’ word knowledge and wanted tools to evaluate how well students could generalise their spelling knowledge to similar words. (These conversations led to the creation of the APSI for data collection in the project.)

Research Question 1

Classroom observations were completed three times (fall, winter and spring) in 14 participating classrooms and twice in 3 classrooms, resulting in 48 observations. Average total Classroom Observation Tool scores for the teachers ranged from 4 to 26, and the average total score was 16.25 out of 30 total points ($SD = 6.78$). Across all the classrooms, several scores were consistently high on the Classroom Observation Tool, for example, preparation and organisational materials ($M = 2.69$, $SD = 0.87$). Other scores were consistently low; for example, while many classrooms had word study notebooks, few used them to write reflections or included teacher feedback ($M = 0.63$, $SD = 1.09$). We discussed the de-identified data in our focus groups and selected five categories of scores for closer analyses because they (1) differed among classrooms and (2) reflected the components of most interest to the teachers. These categories were grouping, teacher talk, student-to-student talk, reflection and student engagement.

Grouping indicated how differentiated the instruction was. A 3 indicated that all children were assigned work based on their assessed developmental spelling level and a 0 indicated that everyone in the class was doing the same work. Classroom scores averaged 2.38 ($SD = 1.20$).

Teacher talk indicated the quantity and quality of teacher talk about the lesson. A 3 indicated that the teacher was observed leading explicit spelling instruction, asking students open-ended, higher-order questions and restating and summarising student ideas. A 0 indicated that teachers talked very little or that the talk was primarily giving instructions or managing behaviour. Classroom scores averaged 2.06 ($SD = 0.93$).

Student-to-student talk indicated the amount of talk between students. A 3 indicated that students were observed having substantial conversations among themselves about the

content of the lessons. A 0 indicated that students were not observed talking to one another or that their talk was not about the content of the lesson. Classroom scores averaged 1.25 ($SD = 1.13$).

Reflection indicated the extent to which the lesson arrived at some conclusions or ‘big ideas’ about the target words. A 3 indicated that teachers were observed guiding conversation to help the students arrive at some conclusions about the overall point of the lesson. A 1 or a 2 indicated that the teacher was observed directly stating the conclusion. A 0 indicated that this lesson did not arrive at a specific conclusion. Classroom scores averaged 1.25 ($SD = 1.06$).

Student engagement reflected the proportion of students who were engaged with the word work during the lesson. A 3 indicated that almost all of the students were observed as engaged and on task during the lesson. A 0 indicated that most students were not engaged or on task during the lesson. Classroom scores averaged 1.88 ($SD = 0.96$).

Classroom organisational models. Through discussions of these scores and the observations, we identified three primary models of organising practice, which reflected the different priorities of the teachers and were reflected in the patterns of Classroom Observation Tool scores (see Figure 1).

Model 1: Differentiated, teacher-led small-group work. Seven teachers organised their students into small groups (5–10 children in each) based on spelling assessment data and conducted instruction through short meetings (10–15 minutes) with those small groups. A typical lesson would be integrated into the literacy block and take about an hour to rotate through all the groups. The small group engaged in sorting of words with target patterns and teacher-led discussion of the words, their spelling and sound patterns, and their meanings. These small-group discussions were carefully differentiated ($M = 3$, $SD = 0$) and typically rich in teacher talk ($M = 2.57$, $SD = 0.53$), but they were teacher-led with limited time for student-to-student talk ($M = 1$, $SD = 0.82$) and many groups never got around to reflection ($M = 1.14$, $SD = 0.90$). Students working directly with the teacher were generally engaged ($M = 1.86$, $SD = 0.90$). Students who were not working directly with the teacher worked independently or in pairs on other literacy activities, for example, illustrating their sort words or writing in their journals.

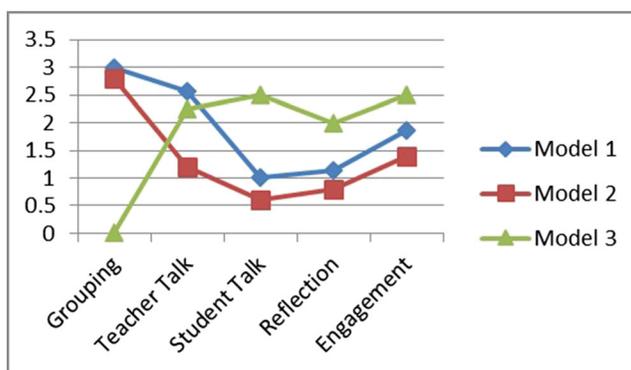


Figure 1. Mean Classroom Observation Tool scores for teachers by organisation type. COT, *Words Their Way* Classroom Observation Tool. [Colour figure can be viewed at wileyonlinelibrary.com]

Model 2: Differentiated, individual work. Six teachers provided students with differentiated materials, resulting in high grouping scores ($M = 2.8$, $SD = 0.45$), but most of the work was conducted independently, with the teacher briefly introducing sorts and giving instructions but spending most of the lesson circulating to supervise, answer questions and manage behaviour. There was little teacher talk, and most of it was devoted to instructions and behaviour management ($M = 1.2$, $SD = 0.84$). Children were not seated with children working with the same materials, so they did not work together, and there was little student-to-student talk ($M = 0.6$, $SD = 0.89$). Because there was little talk in general, there were few opportunities for reflection ($M = 0.8$, $SD = 0.84$). Engagement was generally high at the beginning of the lesson when the activities began but tapered off as some students finished the work before others ($M = 1.4$, $SD = 0.89$). This model was most prominent in classrooms that utilised one-to-one technology to ‘push’ different sorting materials to individual students and allow them to work independently.

Model 3: Whole class instruction. Four teachers chose not to differentiate spelling instruction. Instead, all the children in a class worked on the same content at the same time (grouping $M = 0$, $SD = 0$). Typically, lessons included whole-group discussion of words, followed by small-group work on particular activities with the words, followed by whole-group discussion of what they learned from these activities. The teacher alternated between leading these whole-group discussions and facilitating small-group work (teacher talk $M = 2.25$, $SD = 0.96$). These classrooms were characterised by high levels of student-to-student talk ($M = 2.5$, $SD = 1$), reflection ($M = 2$, $SD = 1.41$) and engagement ($M = 2.5$, $SD = 1$).

Research Question 2

Descriptive statistics and correlations. Means and standard deviations for student scores on the PSI and APSI by grade level are presented in Table 3. Cronbach’s alpha for the PSI was .92 and for the APSI was .91. Agreement between the scorers was 98% for binary scoring and 96% for Levenshtein difference scoring. The PSI and APSI scores were highly and significantly correlated at both the fall ($r = .87$, $p < .01$) and spring ($r = .89$, $p < .01$). PSI scores were significant predictors of APSI scores, explaining 80% of the variance in fall and 81% in the spring ($p \leq .001$).

Table 3. Descriptive statistics and correlations for Primary Spelling Inventory and Alternative Primary Spelling Inventory scores.

	Fall 2017			Spring 2018		
	PSI	APSI	<i>r</i>	PSI	APSI	<i>r</i>
1st grade	5.89 (3.68)	5.56 (3.41)	.85*	12.23 (5.48)	11.17 (5.67)	.85*
2nd grade	12.28 (5.07)	14.47 (4.42)	.62*	18.46 (5.06)	17.41 (4.35)	.86*
3rd grade	16.66 (7.06)	15.38 (5.75)	.92*	20.30 (5.86)	18.43 (5.34)	.91*
Total	10.14 (6.47)	10.52 (6.37)	.89*	14.77 (6.42)	13.65 (6.26)	.90*

APSI, Alternative Primary Spelling Inventory; PSI, Primary Spelling Inventory.

* $p < .01$.

Paired sample t-tests. Paired sample *t*-tests indicated that total scores on the APSI were significantly lower than the PSI in spring but not fall. Students performed significantly differently on 12 out of 26 item pairs analysed in the fall and 12 pairs in the spring. The percentages of students who got each item correct on the two forms are illustrated in Figure 2. A Bonferroni correction was applied to the alpha levels due to the number of contrasts, so only contrasts significant at the $p < .001$ were considered significantly different.

The data indicated that while spelling ability as measured by the PSI and APSI increased over the course of the year and by grade level, the match between the PSI and APSI did not evidently increase with grade level or over the year. When the word-level results were presented to the focus groups, teachers noted that these word pairs included spelling features mastered in each developmental spelling stage. Reviewing their own students' data, they noted the errors on these word pairs often included accurate representations of the phonetic structure that substituted common vowel patterns (*gleem vs gleam*) or did not acknowledge the morphological structure of the word (*frise vs fries*; *stampt vs stamped*). We hypothesised that as students' spelling knowledge develops, they become more aware of phonetically plausible ways to spell a word, and choosing the correct one will depend on the familiarity and semantic difficulty of the word. At our next meeting, the first author presented research on semantic aspects of words that interact with word recognition and spelling, in particular word concreteness and age of acquisition (Brysbaert et al., 2014; Hiebert, 2012; Hiebert et al., 2019; Kuperman et al., 2012).

Multilevel modelling. To model the ability of students to analogise from known to unknown words that share spelling patterns, the first author used multilevel regression to analyse the spring spelling data (Raudenbush & Bryk, 2002). The data were analysed using two forms of APSI word-level scores as outcome variables: binary scores (0 or 1) and Levenshtein distance scores. Four thousand, six hundred and twenty-eight APSI items were nested in 178 students who were nested in 10 classrooms. The interclass correlation coefficients indicated that 12% of the variance in correct spellings was between children and 39% of that between-child variance was between classrooms, and 15% of the variance in Levenshtein distance scores was between children, with 10% of that variance between classrooms. All models were run using full maximum likelihood. Binary variables were

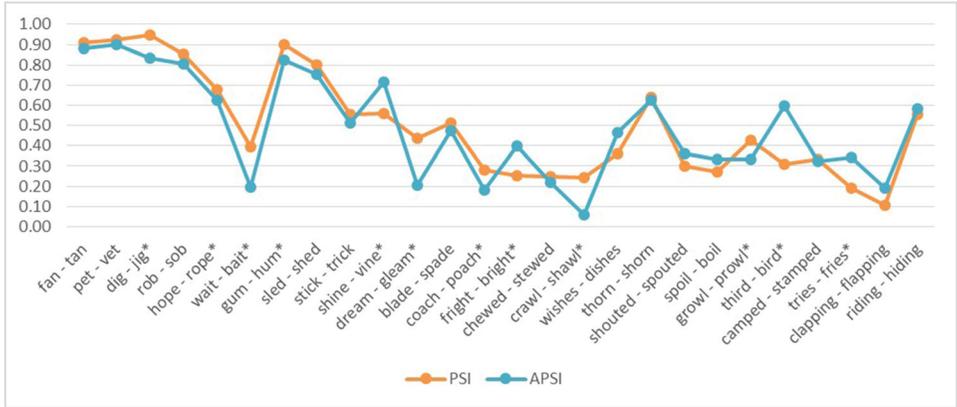


Figure 2. Item performance on the Primary Spelling Inventory and Alternative Primary Spelling Inventory. APSI, Alternative Primary Spelling Inventory; PSI, Primary Spelling Inventory. *indicates that spelling performance on this item pair differed significantly, $p < .001$. [Colour figure can be viewed at wileyonlinelibrary.com]

entered uncentred; continuous variables were entered group (Level 2) or grand (Level 3) mean centred.

At Level 1, the tested predictors were the students' PSI rime-matched item score and measures of word difficulty: word length (in letters), frequency (dispersion), concreteness, age of acquisition and part of speech measure from Word Zone Profiler (Hiebert, 2012). At Level 2 (students), the total score for the PSI was included at the student level to control for overall spelling ability. At Level 3 (teachers), grade level, overall Classroom Observation Tool scores and model of instruction were included. Insignificant predictors were removed to specify the final models.

The results of the final models are presented in Tables 3 and 4. For the logistic model predicting correct spellings on the APSI (Table 4), odds ratios indicated that spelling the paired item on the PSI correctly increased odds of spelling the corresponding word on the APSI correctly by six times, but more complex words (higher age of acquisition, concreteness and word length) were associated with lower odds of spelling the item correctly. Overall total PSI scores were also associated with a greater chance of spelling each item correctly, and each increase in grade level doubled the odds of analogising correctly.

The model predicting Levenshtein difference scores had similar results (Table 5), although it should be noted that the outcome variable represents the number of errors, so the direction of the effects is reversed. The intercept was 1.24, indicating that an average student in the sample who did not spell the matched PSI word correctly would omit, insert or substitute an average of 1.24 letters on a word of average difficulty. Spelling the matched PSI word correctly reduced the average error to 0.56 letters. Higher grade level and overall PSI scores were associated with fewer errors, and more complex words were associated with more errors including the word frequency measure, which was not associated with the binary performance scores. There was no significance effect of part of speech, teachers' Classroom Observation Tool scores or model of instruction in either model.

Table 4. Results of logistic multilevel regression model predicting binary Alternative Primary Spelling Inventory item scores.

	Coefficient	Odds ratio	CI
Intercept	-2.01	0.13	(0.07, 0.27)
Level 1			
PSI item*	1.84	6.30	(5.27, 7.52)
Age of acquisition*	-0.34	0.71	(0.68, 0.75)
Concreteness*	-0.30	0.75	(0.63, 0.87)
Word length*	-0.58	0.56	(0.52, 0.60)
Level 2			
PSI total*	0.12	1.13	(1.10, 1.15)
Level 3			
Grade*	0.69	2.00	(1.30, 3.09)

CI, confidence interval; PSI, Primary Spelling Inventory.

* $p < .001$.

Table 5. Results of multilevel regression model predicting Alternative Primary Spelling Inventory Levenshtein distance scores.

	Coefficient	SE	t-ratio	f^2
Intercept	1.24	0.07	18.40	
Level 1				
PSI item**	-0.68	0.03	-21.22	0.11
Age of acquisition**	0.11	0.01	10.57	0.03
Concreteness*	0.07	0.03	2.60	0.00
Dispersion**	0.08	0.01	-6.29	0.01
Word length*	0.16	0.01	15.71	0.06
Level 2				
PSI total*	-0.04	0.00	-9.48	
Level 3				
Grade*	-0.27*	0.09	-2.82	

PSI, Primary Spelling Inventory.

* $p < .01$.

** $p < .001$.

DISCUSSION

As a research–practice partnership, the Word Nerds project generated data and insights important to policy, theory and practice and laid groundwork for future research (Penuel & Gallagher, 2017). First, the focus groups indicated that state-level policies that mandated particular outcomes left teachers and literacy leaders largely at sea. Second, implementation of a spelling approach that was not a scripted curriculum resulted in varied instructional approaches with different priorities. Finally, student data supported the underlying theoretical models and linguistic approaches to spelling instruction but highlighted the need to connect spelling instruction to vocabulary instruction early in spelling development.

When policy and best practices conflict

The collaborative structure of this project allowed us to explore variety in implementation in spelling instruction through the lens of what teachers do and why. Some teachers prioritised direct instruction (Model 1); others differentiation (Model 2); and others standards-based content and peer collaboration (Model 3). These different approaches offered different affordances and advantages to teachers and students. In Model 1, every student had a teacher-directed and scaffolded spelling lesson at his or her developmental level. However, children who were assessed at lower spelling levels at the beginning of the year rarely had an opportunity to work at grade level. Further, while the teacher worked with one group, the majority of the class needed to work independently on extension activities, and most of each child's instructional time was spent on these activities. Spelling instruction in Model 2 took little time; everyone worked at their own level for 10–20 minutes and then moved on to a new subject. However, it provided little opportunity for scaffolding

either from teachers or peers. Model 3 included opportunities for both teacher and student-to-student talk and engaged all students simultaneously in grade-level content. The teachers who used this model indicated that they did so because they saw the content as necessary for all their students (e.g., short vowel word families in first grade); however, lessons were not tailored to individual students' needs.

As described above, teachers reported feeling conflicted about differentiation, a hallmark of stage-based approaches to spelling (Treiman, 2017) and a central tenet of *Words Their Way* (Bear et al., 2019), and the instructional models influenced how teachers thought about appropriate difficulty. Children working individually were assigned easier material that teachers felt they could complete independently, whereas teachers were more comfortable challenging children in small-group or whole-group instruction. To illustrate, the first author observed a first-grade classroom work with compound words (Model 3) on the same day she observed a different first-grade classroom working independently (Model 2) on various picture sorts of single-syllable words. Furthermore, teachers seemed to think they could only differentiate the content of the lessons (the words that children were studying), and they did not consider differentiating scaffolding or support, such as instructional strategies that make phoneme–grapheme mapping more explicit (e.g., word building; Elkonin boxes). In addition, the relationship between semantic difficulty and spelling difficulty took the teachers by surprise. *Words Their Way* does include lessons that focus on morphemes in compound words, prefixes and suffixes, and Greek and Latin roots; however, most of these lessons are intended for students in the later developmental stages, typical of students in third grade and beyond. The differentiated approach to spelling instruction in the *Words Their Way* programme meant that struggling students would receive little instruction in spelling–meaning connections before the end of third grade.

Thus, while research has established that teachers struggle to implement best practices in spelling, the question is which best practices? The Word Nerds project illustrated a continuum of practices that highlight different evidence-based priorities. Given the wide variation of implementation of the *Words Their Way* approach, identifying these larger patterns of organising instruction may be useful in future research that attempts to measure its effectiveness. Further, it is important for leaders and policy-makers to know that teachers experience some best practices to be in tension with one another and with state-level policy and equity. Overall, the findings of the Word Nerds project indicate that spelling instruction would be improved by articulating clear priorities, providing opportunities to analyse practice and student data collaboratively and offering more instructional strategies to support both orthographic and semantic learning of words.

When theory is practice

While we often view research as either basic (theory-driven) or applied (practice-driven), we found that teachers had substantial *personal theories* about how their students learned that overlapped with *formal theories* extant in the research literature (Ravitch & Riggan, 2016). At the beginning of the project, teachers reported feeling frustrated that their students' general spelling knowledge was not improving. Analysing the data collaboratively allowed teachers to see the bigger picture and recognise their students' successes. In addition, teachers had the opportunity to observe, hypothesise and test the role of semantic difficulty in spelling, with immediate implications for practice. After discussing these findings, teachers in the project described including more meaning support during spelling

lessons; for example, one first-grade teacher taught that *mow* is spelled like *snow*, while miming mowing a lawn.

The students in this project showed evidence of spelling by analogy as predicted by theory and established in prior research. Overall spelling scores on the two assessments were highly correlated and increased together over the course of the year and from grade level to grade level. Knowing how to spell a word with the same rime increased the odds of spelling a word by six times and reduced the average spelling error from more than one letter per word to half a letter per word. Overall, these findings suggest that as children's mental lexicons increase, they have a larger base to use for analogy and phonological recoding, as described by Ehri (2005) and Share (1995), and that rimes are important units for analogy, as described by Goswami (1999) and others. However, the ability to spell rime-matched words correctly was not more pronounced for easier spelling items or for older children. Word frequency indicates the number of opportunities a child has had to recode a word (Perfetti & Hart, 2002; Share, 1995), and while it did not increase the odds of spelling a word correctly, more frequent words had fewer errors per word, indicating high lexical quality. The fact that measures of semantic difficulty decreased the odds of spelling a word correctly suggests that word meanings are invoked to distinguish between plausible spellings for words (e.g., *jig* vs *gig* and *shawl* vs *shawl*). As described by the lexical quality hypothesis (Perfetti & Hart, 2002), a high-quality word representation is one that links the phonological, orthographic and semantic characteristics of a word; however, spelling instruction for the early spelling stages often focuses primarily on the phonological–orthographic links (see Table 1). Unless there is a specific representation for a word in lexical memory, children may generate plausible but not necessarily correct spellings. High-quality word representations may create a direct path between a word's meaning and spelling, in which children rely on a reverse set for variability process to retrieve correct spellings from memory (is *mow* spelled like *go* or *foe* or *snow*?).

Research as partnership: Insights and limitations

In response to a perceived literacy crisis, Michigan districts are currently adopting new early literacy programmes, materials and initiatives, often with little professional buy in from teachers and little research evidence about their effectiveness. Through working with teachers and school leaders, the Word Nerds project identified a high degree of variability in implementation of *Words Their Way*, indicating there is not agreement among teachers on what this approach entails. Instead, we were able to excavate the key questions that teachers had about their own practice and identify patterns of practice. We also developed tools to examine an instructional outcome that matters to both the teachers themselves and to theory – children's ability to use analogy to spell new words. Teachers reported that participating in the project increased their understanding of their own instruction and their students' development. They also felt empowered by the project to adapt the recommendations of *Words Their Way* to meet the needs of their particular students.

The research–practice partnership format also presented challenges and limitations. In the initial stages of a partnership, the goal is to build trust and understand current practices (Coburn & Penuel, 2016; Penuel & Gallagher, 2017). The Word Nerds team was the right size for collaboration and discussion, but too small to observe connections between practice and student outcomes. Experimental and quasi-experimental studies are needed to compare the results of these models on children's spelling gains and to understand which

elements are most crucial for learning. Furthermore, deciding to work in partnership constrained the project to student data districts were willing to share and to areas where teachers and leadership were interested in spelling instruction and had the time and motivation to take on the partnership work, resulting in a sample that was primarily white, affluent and not struggling with literacy.

CONCLUSION

Policies, professional development and official documents and recommendations offer guiding principles to improve literacy instruction but are often agnostic about approach, particularly in the United States where local control of schools is highly valued. In addition, top-down directives are unlikely to change daily teaching practice without including teachers' expertise and experiences. Further, teachers' knowledge and practice improve through opportunities to put knowledge into practice with local support and feedback (e.g., Flynn et al., 2021; Hindman et al., 2020). Teachers joined the Word Nerds project in part for a community of practice and noted that, on their own, they had struggled to implement new practices, analyse student data, and recognise progress and improvement. While the politics of schools in the United States is hyperlocal and mandated curricula are often resisted, many teachers who want to implement the recommendations need specific, daily guidance and materials, particularly when they are trying new approaches that they did not experience in their own schooling or teacher training. Research–practice partnerships can offer scholars the opportunity to understand and analyse questions of implementation and teachers the confidence and desire to continue. Teachers need guidance from policy and research on how to balance differentiation, grade-levels standards and equity. Ongoing professional learning communities, and further studies of those communities, are needed to support teachers who want to make the recommended changes to their instruction.

ACKNOWLEDGEMENT

The research reported in this paper was not supported by external funding.

CONFLICT OF INTEREST

I have conducted professional development for districts using *Words Their Way* materials. I do not endorse or sell the programme but do support teachers in learning to use the materials.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

- Bassok, D., Latham, S., & Korem, A. (2016). Is kindergarten the new first grade? *AERA Open*, 2(1), 1–31. <https://doi.org/10.1177/2332858415616358>
- Bear, D. R., Invernizzi, M. R., Templeton, S., & Johnston, F. (2019). *Words their way: Word study for phonics, vocabulary, and spelling instruction*. Pearson.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteity ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46, 904–911. <https://doi.org/10.3758/s13428-013-0403-5>
- Coburn, C. E., & Penuel, W. R. (2016). Research-practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, 45(1), 48–54. <https://doi.org/10.3102/0013189X16631750>
- Conrad, N. J., Kennedy, K., Saoud, W., Scallion, L., & Hanusiak, L. (2019). Establishing word representations through reading and spelling: Comparing degree of orthographic learning. *Journal of Research in Reading*, 42(1), 162–177. <https://doi.org/10.1111/1467-9817.12256>
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4
- Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A. M. (2012). From spelling pronunciation to lexical access: A second step in word decoding? *Scientific Studies of Reading*, 16(4), 341–359. <https://doi.org/10.1080/10888438.2011.568556>
- Erickson, F. (2014). Scaling down: A modest proposal for practice-based policy research in teaching. *Education Policy Analysis Archives*, 22, 9. <https://doi.org/10.14507/epaa.v22n9.2014>
- Flores, J. G., & Alonso, C. G. (1995). Using focus groups in educational research: Exploring teachers' perspectives on educational change. *Evaluation Review*, 19(1), 84–101. <https://doi.org/10.1177/0193841X9501900104>
- Flynn, N., Powell, D., Stainthorp, R., & Stuart, M. (2021). Training teachers for phonics and early reading: Developing research-informed practice. *Journal of Research in Reading*, 44(2), 301–318. <https://doi.org/10.1111/1467-9817.12336>
- Fresch, M. J. (2007). Teachers' concerns about spelling instruction: A national survey. *Reading Psychology*, 28(4), 301–330. <https://doi.org/10.1080/02702710701545510>
- Fresch, M. J. (2003). A national survey of spelling instruction: Investigating teachers' beliefs and practice. *Journal of Literacy Research*, 35(3), 819–848. https://doi.org/10.1207/s15548430jlr3503_2
- Ganske, K., & Jocius, R. (2013). Small-group word study: Instructional conversations or mini-interrogations? *Language Arts*, 91(1), 23–40. <https://www.jstor.org/stable/24575013>
- Gehsmann, K. M., & Bear, D. R. (2013). *Words Their Way classroom observation tool*. Pearson.
- Gehsmann, K. M., Millwood, K., & Bear, D. R. (2012). *Validating a classroom observation tool for studying developmental word study instruction*. Paper presented at the 62nd annual conference of the Literacy Research Association, San Diego, CA.
- Goswami, U. (1999). Causal connections in beginning reading: The importance of rhyme. *Journal of Research in Reading*, 22(3), 217–240. <https://doi.org/10.1111/1467-9817.00087>
- Graham, S., Morphy, P., Harris, K. R., Fink-Chorzempa, B., Saddler, B., Moran, S., & Mason, L. (2008). Teaching spelling in the primary grades: A national survey of instructional practices and adaptations. *American Educational Research Journal*, 45(3), 796–825. <https://doi.org/10.3102/0002831208319722>
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27(9), 1703–1743. <https://doi.org/10.1007/s11145-014-9517-0>
- Henderson, E. H. (1990). *Teaching spelling*. Houghton Mifflin College Division.
- Henrick, E., Munoz, M. A., & Cobb, P. (2016). A better research-practice partnership. *Phi Delta Kappa*, 98(3), 23–27. <https://doi.org/10.1177/0031721716677258>
- Hiebert, E. H. (2012). *Word Zone Profiler*. TextProject.
- Hiebert, E. H., Scott, J. A., Castaneda, R., & Spichtig, A. (2019). An analysis of the features of words that influence vocabulary difficulty. *Education Sciences*, 9(8), 1–24. <https://doi.org/10.3390/educsci9010008>
- Hindman, A. H., Morrison, F. J., Connor, C. M., & Connor, J. A. (2020). Bringing the science of reading to pre-service elementary teachers: Tools that bridge research and practice. *Reading Research Quarterly*, 55(S1), S197–S206. <https://doi.org/10.1002/rrq.345>
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44, 978–990. <https://doi.org/10.3758/s13428-012-0210-4>
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707–710.

- Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2021). How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development. *Journal of Research in Reading*, 44(1), 10–26. <https://doi.org/10.1111/1467-9817.12313>
- Michigan Department of Education. (2017). www.mischooldata.org
- National Center for Education Statistics. (2016). Digest of Educational Statistics. Retrieved from: <https://nces.ed.gov/programs/digest>
- Ouellette, G., Martin-Chang, S., & Rossi, M. (2017). Learning from our mistakes: Improvements in spelling lead to gains in reading speed. *Scientific Studies of Reading*, 21(4), 350–357. <https://doi.org/10.1080/10888438.2017.1306064>
- Penuel, W. R., & Gallagher, D. J. (2017). *Creating research practice partnerships in education*. Harvard Education Press.
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. T. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy*. John Benjamins Publishing. <https://doi.org/10.1075/swll.11.14per>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Raudenbush, S. W., Bryk, A. S., Cheong, A. S., Fai, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Scientific Software International.
- Ravitch, S. M., & Riggan, M. (2016). *Reason & rigor: How conceptual frameworks guide research* (2nd ed.). Sage Publications.
- Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, 62(10), 1948–1966. <https://doi.org/10.1080/17470210802696104>
- Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11(3), 235–257. <https://doi.org/10.1080/10888430701344306>
- Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, 100(1), 175–191. <http://content.apa.org/journals/edu/100/1/175>, <https://doi.org/10.1037/0022-0663.100.1.175>
- Schlagal, B. (2002). Classroom spelling instruction: History, research, and practice. *Literacy Research and Instruction*, 42(1), 44–57. <https://doi.org/10.1080/19388070209558380>
- Shahar-Yames, D., & Share, D. L. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading*, 31(1), 22–39. <https://doi.org/10.1111/j.1467-9817.2007.00359.x>
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151–218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- Sterbinsky, A. (2007). *Words Their Way spelling inventories: Reliability and validation study*. Center for Research in Educational Policy.
- Treiman, R. (2017). Learning to spell words: Findings, theories, and issues. *Scientific Studies of Reading*, 21(4), 265–276. <https://doi.org/10.1080/10888438.2017.1296449>
- Treiman, R., Hulstlander, J., Olson, R. K., Willcutt, E. G., Byrne, B., & Kessler, B. (2019). The unique role of early spelling in the prediction of later literacy performance. *Scientific Studies of Reading*, 23(5), 437–444. <https://doi.org/10.1080/10888438.2019.1573242>
- Treiman, R., & Kessler, B. (2014). *How children learn to write words*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199907977.001.0001>
- Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107–136. <https://doi.org/10.1037/0096-3445.124.2.107>
- Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, 16(2), 122–140. <https://doi.org/10.1080/10888438.2010.542527>
- Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. Guilford.

Dr Laura Tortorelli is an Assistant Professor of Elementary Literacy in the Department of Teacher Education at Michigan State University. She earned her PhD in Reading Education from the University of Virginia. Her scholarship focuses on how texts, technology and teachers support early reading and writing acquisition in school contexts. Her work has been funded by the National Science Foundation and the International Literacy Association. She has been named the Jeanne Chall Visiting Researcher by the Harvard Graduate School of Education and an Emerging Scholar by the Reading Hall of Fame. She is currently an associate editor of *The Reading Teacher*.

Lori Bruner is a doctoral candidate in the Curriculum, Instruction and Teacher Education programme at Michigan State University. Her research focuses on the role that texts play in supporting young children's early language and literacy acquisition. She was recently awarded the 2021 Jeanne S. Chall Research Fellowship from the International Literacy Association, given to just one doctoral student per year whose dissertation research addresses reading.

Received 2 March 2021; revised version received 6 March 2022.

Address for correspondence: Laura S. Tortorelli, Department of Teacher Education, Michigan State University, 620 Farm Lane, East Lansing, MI 48824, USA. Email: Ltort@msu.edu