A team of first-grade teachers began to question their literacy/phonics program, which used a whole-class format for instruction. The teachers felt that teaching the whole class did not allow them to meet the diverse needs of all their students. One of the teachers introduced the rest of the team to developmental-spelling research (i.e., research on the development of spelling proficiency that builds on linguistic interpretations of students’ unconventional spellings as they change over time) and, specifically, to an approach to phonics and spelling instruction called word study. By using developmental-spelling assessments, the teachers believed they could better meet their students’ needs through differentiated word study in small groups.

The new school year began with qualitative spelling assessments and subsequent grouping plans. Many issues arose quickly. When teachers placed their students along a developmental continuum of spelling features, they saw that many students could bypass earlier features that, according to assessment results, they had already learned. The idea of having a group of students skip over easier features made them anxious about meeting the Reading First criteria for systematic phonics instruction. In their previous whole-class teaching, everyone started in the same place and proceeded systematically through the curriculum. Now the teachers were uncertain about choosing different spelling features for different groups. They found themselves spending more time preparing when they were used to depending upon an already prepared program. Management issues associated with differentiated groupings proved to be challenging. They wondered how teachers could take developmental-spelling research and make it practical for the classroom.

developmental-spelling research has been at the fulcrum of several educational movements in the past few decades, two of them diametrically opposed to each other. On the one hand, discovery of rule-governed phonetic logic at work in children’s invented spellings sparked a writing revolution in the primary classroom (Richgels, 2001). In Morris’s (1989) words, “It is doubtful how far process-writing would have progressed in first and second grade classrooms had not teachers been given some kind of logical assurance that spelling is a developmental process” (pp. ii–iv). Some, unfortunately, took these assurances too far and eliminated spelling instruction entirely, a negligent immoderation to which plummeting National Assessment of Educational Progress (NAEP) scores have been repeatedly attributed (“Bad Spelling,” 1995). Nevertheless, developmental-spelling research survived the
reactionary restraints placed on invented spelling in school districts across the United States and is now coming into its own in the midst of a very different political arena. Word study, an approach to phonics and spelling instruction that has grown out of developmental-spelling research and theory, now appears in the phonics component of every major basal reading series, and developmental-spelling assessments are endorsed in policy-driven workshops across the United States (Moats, 2003). Virtually every teacher’s manual in every major reading series at least nominally suggests word sorts, and even the Texas Reading First materials include word study as a best practice for one of the five pillars of evidence-based instruction (First Grade Teacher Reading Academies, 2002, see p. 89). Publishers of textbooks, instructional materials, computer programs, and other commercial ventures are capitalizing on word study as the panacea for the phonics and spelling mandate, a panacea that circumvents the boring, decontextualized nature of spelling books and the meaningless drill associated with the workbooks deployed in the past. The recent proliferation of educational books, curriculum guides, computer programs, and special themed issues of educational journals on spelling witness the advent of spelling as one of the most critical areas of interest and concern not only among reading and English language arts educators but also among other scholars and professionals (Templeton, 2003).

Despite these influences, the potential impact of developmental-spelling research for classroom instruction has yet to be realized. The fundamental idea that invented spellings provide a diagnostic cue to a student’s current understanding of how written words work, and that instruction can be timed and targeted to this understanding, is still, for the most part, overlooked. Part of this oversight may be attributed to a lack of understanding about the English writing system itself—the orthography and the systematic way in which children acquire it. Equally culpable is the propagation of mixed messages from commercial publishers and from policymakers regarding the nature of systematic instruction.

To teach English orthography, teachers must sift through bits and pieces of information they hear and read regarding effective literacy instruction and separate the wheat from the chaff, especially now with the push for explicit, systematic phonics and spelling instruction. The National Reading Panel (National Institute of Child Health and Human Development, 2000), for example, reported advantages for synthetic phonics (i.e., explicit instruction about how to convert letters to sounds and then how to blend the sounds together) over analytic phonics approaches despite the fact that teachers almost always teach the sounding out strategy along with other word analysis strategies such as chunking word parts. According to Pressley (2002), teachers routinely complement synthetic phonics with the teaching of word families and other orthographic units, and this practice is supported by recent research that shows an advantage of combined approaches in improving spelling and word recognition over single approaches such as synthetic phonics (Berninger et al., 1998; Lovett, Lacerenza, Borden, Frijters, Steinbach, & De Palma, 2000). Furthermore, newer, alternative approaches, such as word sorting and make-a-word, have an increasingly promising research base (Hall, Cunningham, & Cunningham, 1995; Joseph, 2000; Weber & Henderson, 1989). So, how do teachers go about implementing systematic phonics and spelling instruction these days? How can developmental-spelling research help teachers decide what continuum of phonics and spelling features makes the most systematic curriculum? And how can they teach such a curriculum systematically and still meet diverse students’ needs?

A system within a system:
The history of English spelling informs teaching

A systematic curriculum for phonics and spelling must, first and foremost, reflect the structure of the writing system itself. For this reason, early research on spelling focused on chronicling English orthography from a historical perspective (Craigie, 1927; Scruggs, 1974; Vallins, 1973). This early work by language historians described the evolution of English spelling from the first written artifacts found in seventh-century monasteries to modern English as we spell it today. The historical record of English orthography is central to a definition of systematic because the evolution of written English explains why our spelling system is the way it is—how it is organized and the key features within this organization. The historical record reveals a systematic layering of alphabet, pattern, and meaning as these elements were seamed together across time (Henderson & Templeton, 1986). Knowledge of this system better equips teachers to teach their students how English orthography works systematically to represent a balance of sound and meaning (Wolf & Kennedy, 2003).
The alphabet tier was established during the time of Old English, which spanned between the Germanic invasions of England in the sixth century to the conquest of England by William of Normandy in 1066. Old English was an inflected language whose written counterpart enjoyed a remarkable consistency in letter–sound correspondence. The long vowels in Old English were pronounced close to the way they are in modern French and German today. That is, long-\(e\) was pronounced like the modern long-\(a\), long-\(i\) was pronounced like the modern long-\(e\), and long-\(a\) was pronounced like the broad-\(a\) in father. A single vowel stood for both long- and short-vowel pairings. Nevertheless, Old English laid the foundation for the alphabet to systematically represent individual speech sounds. The simplicity and consistency of the grapheme–phoneme correspondence in Old English were such that, armed with a phonetic guide to letters and sounds, modern readers can still read Abbot Aelfric’s Lord’s Prayer similar to the way Aelfric himself may have recited it over a thousand years ago (Henderson, 1990).

Old English is relevant to teachers today because young children spell like little Saxons as they begin to read and write (Henderson, 1981). Armed with only a rudimentary knowledge of the alphabet and letter sounds, neophyte readers use their alphabetic knowledge quite literally. They rely on the sound embedded within the names of the letters to represent the sounds they are trying to represent (Read, 1971). This strategy works quite well for consonants when the names do in fact contain the corresponding speech sounds (Bee, Dee, ef, eS, etc.). It works less well for letters that have more than one sound (C, G), and it doesn’t work at all for consonants with names that do not contain their corresponding speech sounds (W: double you; Y: why; H: aich, etc.). Short-vowel sounds are particularly problematic for novice readers because there is no single letter that “says” the short-vowel sound. As a result, beginning readers choose a vowel whose pronounced name is closest by place of articulation to the targeted short vowel (Beers & Henderson, 1977; Read). The result is similar to Old English, in which a single-vowel stood for two sounds. Today, a beginning reader may spell BED for bad, band, or bead. Rayner, Foorman, Perfetti, Pesetsky, and Seidenberg (2001) asserted that early spelling inventions such as these indicate a child’s grasp of the alphabetic principle and presage the application of this principle in decoding (see p. 41).

The stability of Old English, however, was permanently altered by a massive influx of French words after the Norman Conquest in 1066. Because these words entered the existing language through bilingual Anglo-Norman speakers, some of the French pronunciations were adopted, too. Also, because the scribes who wrote the new words were biliterate, they used the French spelling for the new vocabulary and even applied French orthographic conventions to the spelling of some English words as well. Old English was thus overlaid with the vocabulary and spelling traditions of the ruling class, the Norman French. This complex interaction of pronunciation change on top of the intermingling of French and English spellings led to a proliferation of different vowel sounds represented by various vowel digraphs. The extensive repertoire of long-vowel patterns today is attributable to this period of history, which also accounts for the various pronunciations of the ea vowel pattern in words like bread, thread, great, break, meat, and clean. To be sure, this merging of oral and written traditions disrupted the relatively consistent relationship between letters and sounds that existed in Old English. However, in Middle English, a second tier of order was superimposed on the alphabetic foundation—a pattern tier in which letter combinations reflected roughly the language of origin.

During this time, somewhere around the 14th century, the pronunciation of long vowels gradually changed, and this change is known as the Great Vowel Shift. Although linguists do not know for sure what caused the Great Vowel Shift, it may have been the result of competing stress contours between English and French (Henderson, 1992). (English places stress most often on the first syllable, while in French secondary or tertiary stress is common—English speakers say ROBert while the French say RoBERT.) Regardless, the Great Vowel Shift is important in understanding the systematic pairing of long- and short-vowel sounds in Modern English and how children make sense of them as they learn to read and spell today. For example, beginning readers’ short-vowel substitutions (e.g., jet spelled JAT; bit spelled BET) are explainable by the change in the pronunciation of vowels that occurred during the Great Vowel Shift.

During the Great Vowel Shift, each long vowel changed in pronunciation. The new pronunciations took the place of older pronunciations, forcing the older vowel into the next higher or lower articulatory position. For example, the old long-\(a\), which had been pronounced as a broad-\(a\) as in calm, became the modern long-\(a\) whose sound we hear today in words like wave, name, or take. But in Old English, long-\(e\) had that spot; long-\(e\) was pronounced like a modern
long-a—*mean* was pronounced *main*. This state of affairs forced the long-*e* to move to a different spot, so it appropriated the modern pronunciation of long-*e*. Unfortunately the old long-*i* used to be pronounced like the modern long-*e* (*bide* was pronounced *heed*) so it too had to change. The old sound of long-*i* thus dropped down to the space that *a* (formerly pronounced like a broad-*a*) had left, so the words *bide, fine, or five* became the modern diphthong sound we hear when we say these words today: *bab-eed* (*bide*), *fab-een* (*fine*), and *fab-eev* (*five*). Although the pronunciation of the long vowels changed in this systematic way, the pronunciation of the short vowels remained the same. As a result, the long-*i* and the short-*i* today are pronounced as they are in the pair of words *ripe* and *rip*—whereas in Old English the single letter *i* represented the pair of vowel sounds heard in *read* and *rid* (Henderson, 1990).

It is uncanny that beginning readers must also make the “Great Vowel Shift” to make the transition from a purely linear, alphabetic stance to a more efficient use of patterns. As students’ reading vocabularies expand sufficiently to sustain independent reading, their lexical store of known words leads them away from letter-name short-vowel substitutions to the correct spelling of short vowels and to find other ways of representing the long vowels. Students begin to take heed of additional silent vowels used to indicate a long-vowel sound and begin to “use but confuse” long-vowel patterns much like the Norman French (Invernizzi, Abouzeid, & Gill, 1994). Spellings like *ROOTE BEER TAITS REELY SWETE* (*root beer tastes really sweet*) look remarkably like Middle English.

The inclination to borrow words that started during England’s long bilingual period foreshadowed a second era of massive word borrowing during the first half of the 16th century. At that time, classical vocabularies were acquired as educated men and women rediscovered the culture and knowledge of classical Greece and Rome. The Renaissance required a new, expanded vocabulary to accommodate the explosion in learning that occurred during this time. Classical roots and stems had the potential to meet this demand for meaning. Greek roots could be compounded (e.g., *autograph* and *autobiography*) and the meanings of Latin stems were recognizable despite derivational changes (*spectator, spectacular, and inspect*). So, to the orthographic record of English history was added a third layer of *meaning*, resulting in a complex but orderly system composed of alphabet, pattern, and meaning resulting in Modern English.

The spelling/meaning relations inherent in words brought into English during the Renaissance have important implications for vocabulary instruction today as students move through the intermediate grades and beyond (Templeton, 1991). As students explore how spelling visually preserves the semantic relationships among derivationally related words (e.g., *bomb* and *bombard*), vocabulary and spelling instruction become two sides of the same knowledge coin. Chomsky and Halle (1968) used the word *muscle* to show how the seemingly arbitrary spelling of some words is, in reality, central to understanding the meanings of related words (*muscle, from the Latin* *musculus: muscular, musculature*). The silent C in the word *muscle* represents a morphemic aspect of written English that preserves its etymological history.

The evolution of English orthography through tiers of alphabet, pattern, and meaning provides a historical structure through which to organize a systematic curriculum for phonics and spelling. It also provides an instructionally transparent, harmonious model for learning to read and write (Bear, Invernizzi, Templeton, & Johnston, 2003). Our spelling system is *alphabetic* because it represents the relationship between letters and sounds established by the Anglo-Saxon scribes. We can match letters—sometimes singly, sometimes in pairs—to speech sounds from left to right and create words. This is the goal of beginning readers and writers who must move from partial to full use of alphabetic cues to blend both letter sounds to encode and segment speech sounds to encode. The *pattern* layer overlays the alphabetic layer. When we look beyond single letter–sound matches and search for patterns that guide the groupings of letters, we find consistent categories of patterns that relate to categories of vowel sounds and the language of origin, largely Norman French (Bear, Invernizzi, et al.). The consolidation of letter patterns affords greater fluency in oral reading and writing, and this orthographic advancement heralds the onset of independent reading (Bear, 1992; Chall, 1983; Ehri & McCormick, 1998). In the third layer of English orthography, the meaning layer, groups of letters (prefixes, suffixes, and Greek and Latin stems) represent meaning directly. Students operating within the meaning layer of English orthography have relatively automatic word recognition and encoding skills, and thus their minds are free to think as rapidly as they can read and write (Gough & Hilleneger, 1980). Understanding the synchronous advances in reading, writing, and spelling development is crucial for balanced literacy instruction (Cramer, 1998).
Although most students are introduced to English spelling through the phonics component of a basal reading series, these programs may or may not conform to this general progression from sound to pattern to meaning in their phonics or spelling component. In the basal phonics lesson, the emphasis is usually on grapheme–phoneme correspondences as students learn how letters represent speech sounds. The presentation of vowel patterns and meaning elements (morphemes), the more advanced features in our orthographic history, is less deliberate. As Henry (1989) noted, “Rarely is this information organized for either teachers or their students” (p. 136).

Henry did organize this scope and sequence according to what she referred to as a historical structural perspective, and she examined the effectiveness of instruction in word features organized by the historical structures associated with the Anglo-Saxon, Romance, and Greek languages. Henry labeled these features letter—sound correspondences, syllable patterns, and morphemes, paralleling what Henderson and Templeton (1986) referred to as alphabet, pattern, and meaning. Henry (1989) found that third- and fifth-grade students who were explicitly taught historical structures and the specific features characteristic of those structures made greater gains in word recognition and spelling than did a control group. Not only did the curriculum follow a systematic progression of historical principles based on alphabet, pattern, and meaning, but also the content (orthographic features such as Latin stems like tract, spect, dict, etc.), structure (placement of features within words, such as inspect versus spectator), and process (teaching techniques, such as generating, reading, and spelling other examples) were presented in systematic routines. Henry wrote the following:

In the opening teachers stated the purpose and goals for the lesson and introduced the new concept, pattern, and/or generalization. During the middle segment of the lesson, students discussed word features as they read words fitting a specific pattern, generated new words, and spelled numerous words fitting frequently used patterns. They often contrasted this pattern to other similar patterns and became aware of how word structure influences spelling. They spent little time in isolated drill and workbook practice. The lessons allowed the students to think of each concept and strategy as a problem solving activity. To close the lesson, teachers and students summarized and reflected upon the lesson content, structural patterns, and procedures. Follow-up exercises promoted reinforcement of the concepts. For example, students looked for Greek words in their science texts...or looked for affixes in the evening newspaper. (p. 145)

In this example, historically derived orthographic structures, the organization of the content within each structure, and the consistency of the routines for teaching that content provide the system. Organizing the spelling curriculum according to historical orthographic structures also places the corpus of words to be studied in an evolutionary progression. Anglo-Saxon words, the oldest in English, are among the easiest to read and the most familiar. Words like sun, moon, day, and night are high-frequency “earthly” words that populate easy reading material. Anglo-Saxon words most visibly survive in high-frequency prepositions, pronouns, conjunctions, and auxiliary verbs (e.g., have, was, does) where the pronunciation is now quite different from when English was an inflected language. More difficult Norman French words of one and two syllables—words like chance, chamber, royal, feudal, guard, and conquer—do not typically appear in beginning reading texts; they appear with greater frequency in books suitable for the upper elementary grades. English’s most difficult words—words like captive, circular, calculate, imitate, maximum, cumulus, nucleus, hemisphere, hydraulic, and rhombus—are of Latin and Greek origin and do not regularly appear in student reading selections until the middle grades and beyond. Thus, the history of English spelling explicates the frequency bands of written words as well as the appropriateness of certain kinds of words over others in relation to the learner. This historical structural perspective on word frequency has been borne out by more scientific accounts, from early computer studies of surface regularities in English spelling (Hanna, Hanna, Hodges, & Rudorf, 1966) to more recent taxonomies that considered frequency as well as predictability in the ordering of words along a continuum from easiest to more difficult (Venezky, 1995). Henderson (1990) added to our understanding of the role that frequency plays in learning to spell by considering the frequency of specific orthographic features, patterns, and morphemes as these relate to students’ reading experience (Nelson, 1989). Programs that select words by frequency alone without also considering the frequency of specific orthographic patterns do not capture the consistency and predictability of English spelling.

Whereas the evolution of written English and a historical structural perspective may provide an overarching vision of a systematic spelling curriculum, the spelling development of individual students must be viewed in relation to the historic development of English spelling. By looking at individual spelling errors across an array of words organized by structures governing alphabet, pattern, and meaning, researchers have described an invariant order in which
students acquire the features of English orthography (Ganske, 1999; Henderson, 1990; Invernizzi, 1992; Slagal, 1982; Viise, 1994). The order is displayed in its simplest form under each tier of English orthography in Table 1. The different shadings represent the different stages or phases as described by developmental-spelling researchers. Note that some spelling features may be learned more or less simultaneously and that instruction can influence the order as well; an early emphasis on consonant digraphs and blends, for example, may result in slightly earlier mastery. In addition, later features of one stage may overlap with some of the earlier features of the next stage, as can be seen with prefixes and suffixes at the top of the third column. Nevertheless, the general progressions within and across stages have been replicated in studies with many different groups of students from preschoolers (Templeton & Spivey, 1980) through adults (Bear, Trucx, & Barone, 1989; Worthy & Viise, 1996), as well as across socioeconomic levels and dialects (Cantrell, 2001; Stever, 1980).

For bilingual learners the precise features of this progression are influenced by phonological and orthographic rules of their native language, but the general constructs within the stages appear similar across speakers of alphabetic languages (Gill, 1980; Temple, 1978). For example, research in the English spelling acquisition of Spanish bilingual learners shows that they sometimes use knowledge of their primary language to spell words in their second language (Fashola, Drum, Mayer, & Kang, 1996; Nathenson-Mejia, 1989; Zutell & Allen, 1988). By assessing the orthographic knowledge of Spanish and English, teachers can observe whether the children are applying Spanish rules of phonology and orthography to English or vice versa (Estes & Richards, 2002). Fashola et al. (1996) stressed how important it is for teachers who work with limited-English-proficient Latino students to learn the phonological and orthographic rules of Spanish so they can differentiate rule-governed “transitioning” errors from random errors, and explicitly explain how a given English phonological or orthographic rule is different from the Spanish one (see p. 840). Spelling researchers have already identified which English consonant sounds are problematic for Spanish speakers, and they are currently working toward identifying additional areas that need explicit attention in other alphabetic languages (Bear, Templeton, Helman, & Baren, 2003). For example, Korean students may confuse the liquids /l/ and /l/ in English because in their native language /l/ and /l/ are an allophonic variation; that is, they are treated the same and are represented by the same letter in Hangul (M. Yang, personal communication, September 16, 2003).

### Systematic assessment and systematic instruction

Nothing has been more ubiquitous in U.S. schools than the spelling test on Friday despite the fact that most students do quite well on Friday only to misspell many of the same words on Monday. It is precisely this phenomenon that argues most convincingly against spelling skill as an outcome of rote memorization and suggests, instead, that learning to spell requires the understanding of the phoneme–grapheme regularities, vowel patterns, and morphological conventions that make up a spelling system. Although learning to spell does entail learning specific words, general knowledge is what is needed when
students stumble upon words they have never seen before, or when they try to write words they don’t know how to spell, or when they are not sure of the meaning of a specific word. For example, a student might invoke general knowledge of consonant blends and short vowels to read the word blast even if he or she has never seen the word before. Likewise, general knowledge that words with similar spellings are often related in meaning, such as recite and recitation, may help a student understand the meaning of recital, even if it is unfamiliar (Bear, Invernizzi, et al., 2003). To be sure, specific knowledge of individual words is absolutely necessary to learn to read and spell. Although the word train might be spelled trane, train, or trayne (all are orthographically and phonetically plausible), only specific knowledge will help us remember which one is correct. However, the relationship between specific knowledge and general knowledge of the system is reciprocal. The ability to remember specific words’ spellings is influenced by general knowledge of the system. At the same time, general knowledge of the orthographic system evolves, in part, “from accumulated experiences with specific word spellings” (Ehri, 1992, p. 308). Short of this general understanding, students have no recourse but to use rote memorization for the test on Friday, and the words are easily forgotten by Monday.

Fortunately, spelling researchers have devised valid and reliable means of assessing students’ general knowledge of the orthographic system as well as their specific knowledge of individual words and individual spelling features (cf. Bear, Invernizzi, et al., 2003; Ganske, 2000; Invernizzi, Meier, & Juel, 2003; Schlagal, 1989; Viise, 1994). Qualitative spelling inventories sample student spellings with words that have been carefully selected to reflect the taxonomy of written English. Students’ spellings are typically scored by the presence or absence of specific spelling features in addition to whether the entire word is spelled rightly or wrongly. For example, in assessing student knowledge of consonant blends, train spelled TRANE would receive credit for the tr despite the fact that the whole word is not spelled correctly. Researchers have consistently demonstrated a relationship “power scores” (number of total words correct) and quality of the spelling errors committed (Morris, Nelson, & Perney, 1986; Schlagal, 1989). Students who spell most words correctly tend to commit highly predictable errors of limited variability, while students who spell few words correctly frequently produce unexpected errors with a high degree of variability (Schlagal, 1986). Errors are most interpretable when students misspell neither too few nor too many words. Such a power score indicates students’ instructional level in spelling or where they are in their general understanding of English orthography. The qualitative analyses of the actual spelling errors specify precisely which spelling features students have mastered and which ones must be learned next. Spanish inventories have been developed in a similar fashion for Spanish-speaking students, and these are especially informative when compared to English results for bilingual learners (Bear, Invernizzi, et al.; Estes & Richards, 2002). By methodically assessing students’ orthographic development several times across the year, teachers can ensure that the instruction they plan fits the needs of the students they teach by differentiating spelling instruction in small groups (Bear, Invernizzi, et al.; Ganske, 2000).

Qualitative spelling inventories outline the general terrain of the orthographic system to be learned, and they itemize specific spelling features to be taught systematically in a developmental progression. Spelling researchers have provided evidence for the developmental progression of these features through Guttmann scale analyses. A Guttmann scale is a unidimensional and cumulative scale. For example, a word spelled correctly on a spelling inventory entails spelling less difficult words correctly as well; conversely, misspelling a word on an inventory entails misspelling more difficult words. Bear and Barone (1989), Ganske (1999), Viise (1994), and Invernizzi (1992) have all demonstrated Guttmann coefficients of reproducibility and scalability well within acceptable ranges (reproducibility exceeds .90; scalability exceeds .60), suggesting that qualitative spelling inventories do in fact generate scales that are unidimensional and cumulative. The same cumulative progression has been documented through the use of these inventories with learning-disabled students (Worthy & Invernizzi, 1989), students identified as dyslexic (Sawyer, Wade, & Kim, 1999), and functionally literate adults (Worthy & Viise, 1996). Stage and Wagner (1992), providing additional evidence for the developmental nature of learning to spell, demonstrated the joint influences of phonology and orthography as these interact with working memory across the grades.

Morris explored the concept of a spelling instructional level in a series of studies stemming from a qualitative analysis of error types yielded in student spelling responses to a qualitative spelling inventory (Morris, Blanton, Blanton, Nowacek, & Perney, 1995; Morris, Blanton, Blanton, & Perney, 1995; Morris, Nelson, & Perney, 1986). In the first study, Morris et al. (1986) found positive correlations between overall spelling accuracy (total number of
words correct) and the acquisition of specific orthographic features such as those displayed in Table 1. Next, Morris, Blanton, Blanton, and Perney (1995) followed students in four third-grade and two fifth-grade classrooms where all students received the same instruction at the same pace from the same grade-level spelling book regardless of their prior spelling accuracy. At the end of the year the students were tested on their mastery of curriculum-based words. The top third of each grade could spell most of the curriculum-based word list correctly. However, they found that undifferentiated, whole-group instruction was ineffective for the low-achieving spellers. The bottom third could not spell even half of the words correctly. Finally, Morris, Blanton, Blanton, Nowacek, et al. examined the effects of teaching 48 low-achieving spellers in seven third-grade classrooms at their spelling instructional level instead of their grade level. Half of these low-achieving third-grade spellers were instructed in a second-grade speller (the intervention group) while the other half were instructed in the third-grade speller (control group). Not only did the intervention group score higher on the second-grade posttest, but also they scored nearly as well as the controls on the third-grade curriculum-based posttest, and even higher than the controls on the third-grade transfer test. Taken together, these three studies provide a rationale for establishing an instructional level for spelling instruction through qualitative spelling assessments and for differentiating instruction based on those assessed levels. These findings corroborate those of Schlagal (1982), who found a spread of at least three grade levels in spelling achievement in virtually every class in grades 1 through 6.

How do the basal programs stack up against this developmental evidence and the need for careful assessment to guide instruction? Ever since the National Reading Panel (NRP; National Institute of Child Health and Human Development, 2000) published its report on evidence-based implications for teaching children to read, publishers have beefed up their phonics and spelling instruction within the alphabetic tier to meet the systematic mandate. However, many of the commercial programs do not provide a theoretical or empirical rationale for the scope and sequence of letter-sound instruction (Purcell, 2002). Maslin (2003), in a meta-analysis of first-grade phonics and spelling instruction in five competing basal reading programs (2001–2003 editions), reported correlations ranging from .40 to .96 with the cumulative sequence of orthographic features within the alphabetic tier, as suggested by developmental-spelling research. However, the assessment procedures (mostly in multiple-choice format) did not relate to their instructional sequence. Although all the publishers reviewed paid lip service to using their assessments to differentiate instruction in phonics and spelling, “none of the programs advised starting students either in the kindergarten or the second-grade level program no matter what their strengths and weaknesses were on the assessments” (Maslin, 2003, p. 64). To the contrary, most of the programs suggested that students move through the lessons at the same pace, and suggestions for whole-class instruction were the norm. This is despite the fact that the NRP report clearly stated that “teachers should be able to assess the needs of individual students and tailor instruction to meet specific needs” (p. 11).

Systemizing the system: Translating developmental-spelling research to classroom practice

Spelling research makes clear that the first step toward systemizing the teaching of phonics and spelling is for teachers, as Gentry and Henderson (1980) argued, “to respond appropriately to non-standard spelling” (p. 116). Qualitative spelling assessments and feature analyses assist teachers in determining the instructional levels of their students, and this information allows knowledgeable teachers to place their students within a developmental continuum of systematic instruction. The second step charges educators to differentiate their phonics and spelling instruction according to their students’ instructional spelling levels and, therefore, to implement small-group instruction in their classrooms.

The practice of teacher-directed small-group instruction for reading is not a new phenomenon in the elementary school arena and is supported by research of effective teachers (Taylor, Pearson, Clark, & Walpole, 2000). However, use of group configurations (e.g., circle, seat, center, or four blocks) to manage differentiated instruction in phonics and spelling appears to be slim to nil (Johnston, 2001). Moody, Schum, Fischer, and Jean-Francois (1999) investigated grouping plans suggested in commercial programs and found that grouping plans other than whole group (e.g., small groups, partner dyads, individualized) were suggested only as student-led extensions of an assignment previously taught in the whole group—not as teacher-directed lessons in
small-group formats. This finding is similar to Maslin’s (2003) discovery that the small-group instruction suggested in basal series was recommended only for the reteaching of English-language-learner (ELL) lessons that addressed previously taught skills from the whole-class work.

If we recognize the importance of differentiated instruction, it follows that teachers must adapt the programs that they may be using to reflect effective practice supported by research. Several recent studies emphasize the importance of pacing instruction to students’ instructional levels and engaging them in a search for logical patterns. Juel and Minden-Cupp (2000) found that differentiated phonics instruction was especially beneficial for students with the lowest levels of entering literacy skill. Foorman and Torgesen (2001) also concluded that grouping students is one of the critical elements of instruction that promotes literacy success for at-risk students (see p. 209). Nevertheless, commercial programs directly influence daily teacher practice (Chall, 1996). Because of the ease of implementation, many teachers will want their students to continue through their grade-level program regardless of their instructional level. However, as Henderson (1990) cautioned, teachers must “recognize the absolute limit beyond which no such placement can be permitted” (p. 205).

Once qualitative spelling assessments and feature analyses are completed, teachers should place children in achievement-based groups based on their students’ instructional levels and specific areas of need (Bear & Barone, 1989; Bear, Invernizzi, et al., 2003; Henderson, 1990). These groups are not the static groups of the past (i.e., tracking). Rather, these groups are flexible and dynamic. Teachers regroup as necessary to best meet the changing needs of the students. Effective classroom management, a characteristic of effective teachers, is crucial for grouping plans to be beneficial (Pressley et al., 2001).

Research in the area of professional development and teacher change provides some insight on the challenges related to implementing differentiated word study in the classroom. Johnston (2001) found that although teachers were largely dissatisfied with their students’ spelling skill, they appeared to lack the knowledge and resources needed to teach spelling more effectively. Gill and Scharer (1996) provided a series of six professional development sessions for elementary school teachers attempting to implement developmental-spelling instruction in their school and found that the management of small groups was their overriding concern. In addition to general management concerns (e.g., group rotations), teachers were apprehensive about the meaningfulness of the independent work that students would be doing while not with a teacher-directed group. Some progress was made, however, as teachers learned about English orthography through facilitated discussions about their students’ spelling errors on the qualitative spelling assessments, their change over time, and their response to differentiated word study instruction within a balanced reading and writing curriculum.

Vygotsky theorized that teachers must not only know the developmental level of their students but also work to understand the relationship between their students’ development and the possibilities of instruction (Hedegaard, 1990). Thus, establishing consistent daily and weekly routines is also essential to establishing a systematic program for phonics and spelling instruction. Developmental-spelling researchers favor daily routines that involve sorting or categorizing words in isolation according to similarities and differences in sound, pattern, or meaning and then hunting for other words with similar properties in connected text (cf. Bear, Invernizzi, et al., 2003; Ganske, 2000). Word sorting and word hunting require students to examine the orthographic properties of words in relation to the orthographic characteristics of other words they already know.

Advocates of word study claim that the process of comparing and contrasting orthographic features not only teaches the spelling of specific words but also encourages students to make generalizations about the spelling consistency of other words within a given category (Bear, Invernizzi, et al., 2003; Ganske, 2000). Through word sorting, students note specific letter–phoneme correspondence, patterns, or morphemes (depending on the instructional zone) that generalize across many English words. Although word study promotes an analytic approach to words that can serve students in encounters with unknown words, it is also synthetic in that teachers direct students to deconstruct their words by sound, pattern, or meaning in the process of sorting, depending on where students are on the developmental continuum.

Research on the effectiveness of word sorting is relatively new, but existing evidence indicates that this approach is at least as effective as creating word boxes (boxes for synthetic phoneme segmentation similar to Elkonin boxes) and more effective than more traditional approaches, at least in first grade (Joseph, 2000). Weber and Henderson (1989) studied the effects of computer-assisted word sorting using a random assignment, control group design. The word-sort condition yielded significantly greater gains in word
recognition and contextual reading, and these gains transferred to a standardized spelling test. Hall et al. (1995) and Zutell (1998) provided additional evidence that word sorting is an effective approach to phonics and spelling instruction across the elementary grades. Henry (1989) used a compare-and-contrast approach in her historical–structural curriculum with intermediate students who compared targeted patterns with other similar patterns and reflected on how word structure influences spelling. These students made significant gains in spelling achievement as well as decoding, which Henry attributed to the shared orthography involved with both tasks.

Henderson (1981) devised the concept of word sorting because he was convinced that understanding how children learned to spell words could also provide insight on how they read them. He believed that children’s growing word knowledge encompassed phonological, orthographic, and semantic information and that categorizing written words enabled them to sort out the relations among these linguistic sources. Henderson’s instructional approaches (e.g., word sorts, word hunts, writing sorts) were shaped by his belief that both phonological and orthographic aspects of written words were critical facets in learning to read and write. His work, and the work of his colleagues and students, demonstrated that written word knowledge is developmental and advances progressively in relation to cognitive development, exposure to print, and instruction.

Henderson’s intuitions have since been supported by a number of correlational and longitudinal studies that have consistently identified spelling as an independent contributor to reading acquisition (Cataldo & Ellis, 1988; Ehri & Wilce, 1987; Morris & Perney, 1984). Significant correlations between spelling and various measures of word identification and decoding have also been reported. For example, Ehri (2000) reviewed six correlational studies in which students of various ages (first grade through college) were asked to read and spell words and reported correlations ranging from .68 to .86. In other studies, spelling measures have accounted for as much as 40% to 60% of the variance in oral reading measures (Zutell, 1992; Zutell & Rasinski, 1989). In a two-year study following students from first through third grade, Ellis and Cataldo (1992) reported spelling to be the most consistent predictor of reading achievement. Even intervention studies exploring the added value of supplemental spelling instruction have repeatedly found that students who receive additional spelling instruction perform better on reading tasks such as oral reading, silent reading comprehension, and other reading-related measures in addition to spelling (Berninger et al., 1998; Goulardris, 1992; Graham, Harris, & Chorzempa, 2002; McCandliss, Beck, Sandak, & Perfetti, 2003).

In our statewide data, spelling scores generated from a qualitative spelling inventory administered to 68,817 first graders in the fall of first grade significantly correlated with reading achievement at the end of the year as measured by both word recognition in isolation and oral reading accuracy (Invernizzi, Landrum, Robey, & Moon, 2003). Correlations between spelling scores and reading levels obtained at the same point in time are even stronger, ranging from .79 (end of first grade) to .75 (end of second grade) to .69 (end of third grade). In addition, significant differences are obtained between reading levels, not only for overall spelling score but also for specific orthographic features mastered. For example, at the end of first grade there are significant differences in the inclusion of the silent -e to mark the long-vowel sound in words like slide, brave, or shade, across all reading levels assessed, preprimer through third grade, and these differences are reflected in the increase of the mean score for that particular feature across reading levels. These findings echo those of Stage and Wagner (1992), who found significant differences in phonetically acceptable spellings of nonwords by children in kindergarten through third grade. Results such as these make concrete the theoretical relationship between reading and spelling, underscore the necessity for differentiated instruction, and suggest implications for curriculum developers and publishers.

It takes time for students to develop spelling proficiency. First they must learn the alphabetic principle to generate phonetically plausible spellings. Ehri (1989) asserted that it is precisely this kind of alphabetic knowledge, knowledge of phoneme–grapheme relations, that allows students to make phonetic sense of words and remember them. As knowledge of the phoneme–grapheme relations expands, students begin to acquire many sight words, and they move into independent reading about the same time they begin to navigate through the pattern tier of English spelling. The more students read, write, and learn about the spelling system, the larger their vocabulary grows. Before they know it, students are learning and remembering so many words that their lexicons abound with words of more than one syllable, including, eventually, words of Greek and Latin derivation, words acquired in the meaning tier of written English.

Conclusion

The American Heritage Dictionary (www.yourdictionary.com/ahd) defines systematic as ‘of, character-
ized by, based on, or constituting a system; of or relating for classification or taxonomy; carried on using step-by-step procedures; and, purposefully regular.” In this review of how developmental-spelling research and theory inform classroom practice, we have described the three tiers of English spelling (alphabet, pattern, and meaning) as the system to guide teacher instruction. Further, we have described classroom procedures for implementing a systematic instructional program to include a qualitative assessment of student spellings, differentiated instruction through small-group instruction, and consistent, daily instructional routines. Last, we have provided evidence that students’ knowledge of the orthography within these three tiers is a crucial component of their literacy achievement.

Knowing the system to be taught, devising a system to teach it, and situating word-level instruction within the larger context of comprehensive literacy instruction still leaves the efficacy of specific instructional approaches in question. Continued research investigating the effectiveness of the word study approach for phonics, spelling, and vocabulary instruction is needed. Building on past research, word sorting should be further investigated as an effective approach across the grades as well as an approach to be used in combination with other approaches, such as synthetic phonics. Drawing from and combining with the strengths of spelling approaches such as Orton-Gillingham and word sorting, teachers may better meet the needs of diverse learners. Educators need to know the degree of explicitness necessary to produce significant gains with all different types of learners (e.g., learning-disabled students, ELL students, average-achieving students, gifted students) as well as students at different levels of literacy achievement (e.g., emergent readers, beginning readers, intermediate readers). Contrary to what commercial programs would have us believe, systematic, explicit instruction is not synonymous with everyone on the same page at the same time in the same workbook. Though we already know that whole-class instruction is ineffective, future research must demonstrate the feasibility of differentiating instruction through management plans that utilize small-group configurations. In addition, research is needed to support educators who strive to use teacher-driven techniques in their classroom word study instruction by demonstrating the effectiveness of professional development in developmental-spelling instruction and by providing evidence that knowledgeable teachers are the key to successful classrooms. Last, more research is needed to demonstrate the effectiveness of instructional materials other than those provided by basal reading companies where, more often than not, one size is force-fitted to all.

Although the full potential of alternative approaches to phonics and spelling instruction has yet to be realized, contemporary research will most certainly inform the definition and quality of balanced literacy instruction. The age-old complaint about the boring, decontextualized nature of phonics instruction will likely be supplanted by expressions of surprise and interest in the engaging, inquisitive nature of word study that, in turn, leads to more explicit understanding of English orthography.

MARCIA INVERNIZZI is a reading professor in the Department of Curriculum, Instruction, and Special Education at the University of Virginia, where she teaches courses in word study and diagnosis and remediation, and seminars in reading research. Her research interests include developmental spelling, assessment, and intervention. She is the coauthor of the Phonological Awareness Literacy Screener, the Book Buddled tutoring manual, and Words Their Way. She can be contacted at the McGuffey Reading Center, 405 Emmett Street South, PO Box 400266, Charlottesville, VA 22904-4266, USA, or by e-mail at mail@virginia.edu.

LATISHA HAYES is a reading doctoral student in the Department of Curriculum, Instruction, and Special Education at the University of Virginia. Her dissertation research focuses on the effects of word study in beginning reading instruction and is supported by the 2003 Jeanne S. Chall Research Fellowship. She can be contacted at the McGuffey Reading Center, 405 Emmett Street South, PO Box 400266, Charlottesville, VA 22904-4266, USA, or by e-mail at lhh6e@virginia.edu.

REFERENCES
CANTRELL, R.J. (2001). Exploring the relationship between dialect
and spelling for specific vocalic features in Appalachian first-grade children. Linguistics and Education, 12, 1–23.


FIRST GRADE TEACHER READING ACADEMIES. (2002). Austin, TX: University of Texas System & Texas Education Agency.


Charlottesville, VA: University Printing Services.


