Applied Psycholinguistics **19** (1998), 669–684
Printed in the United States of America

Spelling skills of children in whole language and phonics classrooms

MAGGIE BRUCK
McGill University

REBECCA TREIMAN
Wayne State University

MARKETA CARAVOLAS
McGill University

FRED GENESEE
University of California, Davis

MARIE CASSAR
Wayne State University

ADDRESS FOR CORRESPONDENCE
Maggie Bruck, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1, Canada. Email: Bruck@hebb.psych.mcgill.ca

ABSTRACT
The spelling skills of grade 3 children who had received whole language instruction since they began to learn to read were compared with those of grade 3 children attending a phonics program. The children were asked to spell a list of words and nonwords. Overall, the phonics group produced more accurate word spellings than the whole language group. In addition, the phonics children's spellings of nonwords included more conventional, phonologically accurate patterns. There were no group differences on measures of orthographic acceptability or on preservation of skeletal structure.

Whole language approaches to reading and writing instruction have proliferated in recent years. However, there are few methodologically rigorous evaluations of the risks and benefits of such programs. For example, Stahl and Miller (1989) found only 46 studies comparing the performance of children in whole language classrooms and those in classrooms that used more traditional basal reader instructional approaches. Their review of these studies provided little solid support for the widespread adoption of whole language programs. Stahl and Miller concluded that there was no overall difference between whole language and basal reader programs, that whole language approaches produced somewhat lower effects on comprehension in more recent studies, and that whole language approaches were more effective in kindergarten than in first grade. In a second

comprehensive review of the literature, Stahl and his coauthors (Stahl, McKenna, & Pagnucco, 1994) found only 28 new published studies and 17 new dissertation abstracts comparing the reading skills of children taught in whole language classrooms and those of children taught according to more traditional approaches; only 14 of these studies had sufficient numerical data to allow their inclusion in a meta-analysis. Of particular interest was the focus of these studies: only 20 (44%) included any measures of reading achievement, whereas 49% focused on affective factors. As before, Stahl et al. found that whole language approaches did not differ significantly from more traditional approaches in terms of achievement test performance. However, whole language programs were associated with improved attitudes toward reading.

Even less is known about the influence of whole language programs on spelling and writing. Graham and Harris (1994) reviewed the literature and found 14 studies that examined the effects of whole language on learning to write. The only reliable difference between children in whole language programs and those in more traditional types of programs involved the students' thinking about writing. Students in whole language classes held a meaning-based view of writing, whereas those in conventional classes viewed writing from a skills perspective.

Although the evaluation studies do not provide much support for the "benefits" of whole language programs, neither do they indicate that whole language programs result in lower levels of performance. However, a number of more controlled studies point to certain limitations of whole language curricula. These studies suggest that, if some of the components of more structured phonics curricula were included in whole language programs, children would make greater progress in basic reading and spelling skills. For example, Castle, Riach, and Nicholson (1994) and Rubin and Eberhardt (1996) found that kindergarten children in whole language programs who were given extra instruction in segmentation and who were told to pay attention to the sound structure of words improved in reading and spelling skills (see Uhry & Shephard, 1993, for similar results with grade 1 children). Iverson and Tunmer (1992) found that Reading Recovery programs (a form of whole language early intervention) that included an emphasis on phonological recoding skills resulted in more improvement than Reading Recovery programs that were not supplemented in this way. In a recent study, Eldredge and Baird (1996) compared the writing samples of two groups of grade 1 children. The first group had followed a traditional whole language program from the beginning of kindergarten; they wrote stories as a group (with the help of teachers when necessary) and then were encouraged to try writing on their own. The second group was taught phonemic awareness and phonics in kindergarten and in the first three months of grade 1. From the beginning of November of their grade 1 year, both groups of children were treated similarly: both classes focused on the process of writing (prewriting, drafting, revising, editing, and sharing). In February of grade 1, the children's writing samples were analyzed. Children in the structured program produced more words, more different words, more difficult words, and more correctly spelled words than children in the traditional whole language program. Furthermore, the children in the structured program produced more idea units and wrote better overall compositions than their whole language peers.

Although the latter studies, which examine the effects of adding phonics instruction to whole language programs, provide support for a literacy program that includes a systematic phonics approach, additional data are required to evaluate more fully the benefits and risks of whole language and phonics programs. In the present study, we examine the spelling skills of children who have been educated in a whole language program or in a phonics program. Our study is designed to address four shortcomings of existing studies. First, there are few studies that directly examine spelling. Most of the studies of writing skills do not examine the accuracy of children's spellings but rather focus on the quality of the content or the structure of their written prose passages. Although composition skills are certainly important, automatic and accurate spelling is also an important component of writing skills. A second limitation of the existing studies is that most, particularly those that focus on writing and spelling, only include kindergarten and grade 1 children. As Graham and Harris (1994) noted in their review of the writing literature, there were not enough data to draw conclusions about the writing progress of older children. Such data are crucial in order to examine the long-term effects of different types of educational programs. Even if there are few, if any, differences between instructional programs at the end of kindergarten or grade 1, such differences may occur several years later. Conversely, any advantages detected in kindergarten or grade 1 may be temporary and no longer observable in later grades. A third shortcoming of some existing studies involves the lax criteria for classifying children into whole language or phonics groups. This lax classification procedure is particularly problematic for the study of children past their grade 1 year. In such studies, it is important to ensure that children have in fact fully participated in the program of which they are a member. The data may not accurately portray the effects of whole language programs if some children in the whole language group had previously participated in a phonics program or, conversely, if some children in the phonics group had previously participated in a whole language program. This confound may explain the failure to find significant group differences on some measures where differences would be predicted. For example, Varble (1990) reported that there were no differences in the mechanical aspects of writing (as measured by one global rating for correct spelling, punctuation, grammar, and variety of sentence structure) for children in whole language and traditional classrooms in both grade 2 and grade 6. However, the author conceded that she did not know the previous educational history of the children in her study. It is possible that the children in the whole language group did better than expected because of previous exposure to phonics in earlier years or that the children in phonics programs did worse than expected because some had attended whole language classrooms in earlier years. A fourth limitation of existing studies is that they have mostly examined spelling in terms of "accuracy." We go beyond this one measure by examining a number of different features of children's spellings. This approach was motivated by the view that even errors might reflect a good deal of spelling skill.

The present study examined the spelling skills of grade 3 children who had participated in a whole language or a phonics program throughout their three years of formal literacy training. Toward the end of their grade 3 year, the

children were asked to spell a list of words and a list of nonwords. The misspellings of words and the spellings of nonwords were categorized in terms of their phonological and orthographic acceptability. We also examined the degree to which the spellings preserved the phonological skeleton of the spoken word.

Our analyses were motivated by a body of research that suggests that spelling is a phonological task that involves the ability to segment the target word into its component phonemes and then to assign graphemes (letters or groups of letters) to these phonological units (e.g., Read, 1975; Treiman, 1993). The phonological skeleton measure was devised to assess the first of these skills – the phonemic segmentation of words. If a misspelling contains the same sequence of consonant and vowel graphemes (i.e., the same consonant–vowel structure) as the oral production of the item, then it preserves the phonological skeleton. Thus, the misspelling "machac" for *magic* accurately reflects the consonant–vowel structure of the word, whereas the misspelling "magisk" does not.

The phonological accuracy of children's misspellings has been used as a measure of the degree to which children use information about conventional phoneme–grapheme correspondences in their spelling. For example, the misspelling "beaf" for *beef* is considered to be phonologically accurate, whereas the misspelling "bafe" would be categorized as phonologically inaccurate. Such analyses reflect the degree to which children use phoneme–grapheme information to spell words that they do not know. Children's spellings of nonwords have been used as a measure of their knowledge of sound–spelling correspondences, especially when spelling unfamiliar words.

Knowledge about permissible sequences of letters (sometimes termed orthographic knowledge or visual skills) also plays an important role in spelling. Thus, for example, the grapheme *ck* never occurs at the beginning of English words. A spelling such as "ckat" looks wrong, although one could sensibly generate its pronunciation. For this reason, we also measured the orthographic acceptability of children's word misspellings and nonword spellings.

In English, selection of the correct spelling pattern often depends on the position of the grapheme within the word. This intersection of phonological and orthographic information in the selection of a spelling pattern is referred to as positional constraints. For example, the spelling "papper" for *paper* would make sense if one assigns a sound to each letter. However, if one considers the orthographic surrounds of the vowel *a*, then this spelling cannot be considered appropriate. Thus, as this example shows, skilled reading and spelling require more than the simple assignment of letters to sounds. In previous work, Bruck and Waters (1988) showed that assessments of the phonological accuracy of children's misspellings could change dramatically depending on whether positional constraints were considered. When positional constraints were considered, good spellers' misspellings were phonologically more accurate than those of poor spellers. However, group differences in phonological accuracy diminished when positional constraints were not considered. Although Bruck and Waters did not provide guidelines for the use of one set of criteria versus the other, phonological accuracy that does include positional constraints may more accurately reflect competent spelling skills. This hypothesis was tested in the present study.

We predicted that children in phonics classrooms would be better spellers

overall than children in whole language classrooms, and that the phonics children's better knowledge of phoneme–grapheme correspondences (as measured by performance on a nonword spelling test) would account for their superior performance. This hypothesis is based on the fact that spelling is not taught systematically in whole language classrooms. Instead, the primary emphasis is on getting children to write freely and to produce invented spellings, spelling words according to their own intuitions. If help is required, teachers may sometimes tell children the correct spellings, encourage them to look up the word in a dictionary, or tell them to ask their peers (e.g., see Butler, 1994; Edwards, 1985; Wilde, 1989; Zarry, 1991). Some whole language teachers may show the child or the class how words share a particular spelling pattern. But these short lessons do not systematically sample the spelling patterns of English, nor do they occur on a regular basis. We have observed that, when children in whole language classrooms ask the teacher how to spell a word, the teacher may ask them to sound it out without providing any guidance or suggesting what to do once this has been accomplished.

We also predicted that there would be no differences between the whole language and phonics groups in terms of orthographic acceptability of spellings or preservation of the phonological skeleton of spoken items. Because children in whole language classrooms are encouraged to invent spellings, this might sharpen their awareness of the internal structure of words and their awareness that graphemes are used to represent each segment (see, e.g., Adams, 1990). This experience might allow whole language children to have the same level of insight into the alphabetic principle as children in more traditional programs who are not encouraged to use invented spelling. Finally, we reasoned that, because children in whole language programs are exposed to a wide range of written materials, they may develop an appreciation for the orthographic sequences in English. In more traditional phonics classrooms, children are sometimes explicitly taught about permissible and impermissible orthographic patterns. For example, they may be taught that only the letters *i* and *e* can follow *k*.

METHODS

Participants

We located 54 grade 3 children who, since kindergarten, had attended public schools that strictly adhered to a whole language curriculum. Children were selected from six different public schools in the same school district. They were given reading and spelling tests at the end of their grade 3 year (late May, early June). The average age of these children was 9;2.

The whole language children attended schools that were committed to this approach. From our observations and discussions with their teachers (one of whom is the third author of this article), the children had never had any systematic phonics or spelling instruction. From kindergarten on, these children did a large amount of writing. The children were encouraged to express their thoughts and experiences through writing, using invented spelling. Teachers in the district attended several yearly sessions on how to promote whole language practices in

their classrooms and were actively discouraged from teaching isolated skills. For example, there were no spelling or phonics books in these schools.

The second group of children, the phonics group, included 22 children who attended a private parochial school that adhered to a structured phonics method. These children were given the reading and spelling tests in early March of their school year. The average age of these children was 8;9.

All but one of the phonics children had been enrolled in this school from the beginning of their grade 1 year; the remaining child was enrolled in the school in grade 2. Thus, the children in this group had followed a structured phonics program since the beginning of their formal literacy training and continued to do so even in the grade 3 year. In this school, the children had a phonics workbook, a spelling work book, and a reading series. Phonics was taught as a separate lesson and in conjunction with reading and spelling. Thus, each reading and spelling lesson stressed a particular sound or spelling unit. The children had weekly spelling tests and practiced spelling the words in class by writing these in novel sentences.

Procedures

The children spelled 25 words that sampled a wide range of spelling patterns and orthographic structures. There were 13 one-syllable words, 6 two-syllable words, and 6 three-syllable words. On average, the one-syllable words contained 3 phonemes, the two-syllable words contained 5 phonemes, and the three-syllable words contained 8 phonemes. None of the words contained exceptional phoneme–grapheme patterns or unusual orthographic patterns. The words were chosen to be familiar to grade 3 children. According to Carroll, Davies, and Richman (1971), the average word frequency of the spelling items was 894.

The children also spelled 25 nonwords which were generated by changing one or two of the phonemes (usually the first consonant) of the words. Thus, the nonwords had similar phonological and orthographic structures as the words. The word and nonword stimuli are presented in the Appendix.

We assessed word recognition by means of the Woodcock Reading Mastery Test–Revised (Woodcock, 1987) to determine if the two groups were comparable on this skill. The word recognition subtest of the Woodcock Reading Mastery Test–Revised was administered individually. Within the next week, the children were given the spelling tests in groups of two to four children. They spelled the words and nonwords on two different days. The word spelling task was always given first. For the word spelling test, each word was pronounced and then followed by a sentence containing the word. The word was then pronounced again. For the nonword spelling test, each stimulus was pronounced three times.

For the words, the number of correct spellings was counted. Next, the percentage of spelling errors that were phonologically accurate was counted. Using the scoring system devised by Bruck and Waters (1988), phonologically accurate errors were categorized as either constrained or unconstrained. (Throughout the article, the terms "constrained" and "unconstrained" are used as shorthand versions of the phrases "phonologically accurate according to the constrained sys-

tem" and "phonologically accurate according to the unconstrained system.") A misspelling was categorized as constrained if it could be pronounced like the target word by application of grapheme–phoneme correspondences that included positional constraints (i.e., a certain grapheme is related to a specific pronunciation by virtue of its physical position in the word and the surrounding graphemes). A misspelling was categorized as unconstrained if each phoneme in the word was represented by a grapheme with the corresponding pronunciation. Letter-name spellings were judged to be accurate by the unconstrained system even though these errors might omit a vowel (e.g., "recve" for *receive*; "tnt" for *tent*). Thus, each misspelling was classified as constrained, unconstrained, or phonologically inaccurate. Note that constrained or unconstrained misspellings were not necessarily orthographically acceptable.

For the following pairs of misspellings, the first was classified as constrained, and the second was classified as unconstrained.

	Constrained	Unconstrained
pencil	"pensill"	"pencle"
hard	"harrd"	"hrde"
voice	"voyce"	"vois"
paper	"paiper"	"papper"
draw	"dra"	"dro"

"Pencle" is categorized as unconstrained rather than constrained because the spelling pattern *cle* is usually associated with phoneme /k/ (as in *uncle*) rather than the phoneme /s/. However, if each letter is considered in isolation, then the letter *c* is sometimes associated with the phoneme /s/ (as in *city*). "Harrd" is an example of an orthographically illegal but phonologically acceptable misspelling according to the constrained system. That is, even though there are no words in English that contain the sequence "rrd," there is only one possible pronunciation that could be assigned to this string of letters.

The phonological skeleton of the misspelling was coded next. Misspellings that preserve the phonological skeleton of the spoken word contain the appropriate number of graphemes in the appropriate positions to represent the vowels and consonants in the target word. However, graphemes do not have to correspond to the appropriate phoneme. Examples of errors that preserve the phonological skeleton of a word are "tramatik" for *dramatic* and "affrid" for *afraid*. Errors that do not preserve the phonological skeleton of a word involve the omission or addition of a grapheme. Examples of such errors are "magi" for *magic* and "perotectiv" for *protective*.

The orthographic acceptability of each misspelling was also coded. Regardless of its phonological acceptability or the accuracy of its phonological structure, spellings could be either orthographically acceptable (i.e., the sequence of letters is permissible in English words) or unacceptable (i.e., the misspelling contains sequences of letters that are illegal or nonoccurring in English words). The following are examples of misspellings that were coded as orthographically unacceptable: "receiv" for *receive*; "finlly" for *finally*; "pencl" for *pencil*; "jrmatick" for *dramatic*; and "magik" for *magic*. For example, "receiv" is orthographically illegal because no English words end in *v* (Venesky, 1980); the letter

e is added in such cases. The misspelling "magik" is illegal because the grapheme *k* is generally transformed to the grapheme *ck* when it follows a single vowel letter.

Finally, we examined the frequency of three specific types of errors. The first type was misspellings of postvocalic nasal clusters. Four words contained postvocalic nasal clusters (***pencil***, ***tent***, ***round***, ***transparent***). The second type was misspellings of other initial and final consonant clusters. Eleven words contained consonant clusters in either syllable-initial or syllable-final positions (***hard***, ***drive***, ***draw***, ***player***, ***field***, ***cleverness***, ***surprising***, ***protective***, ***transparent***, ***dramatic***, a***fraid***). The third type was misspellings that contained consonant feature errors. For this category, the confusion of three different phonological features were counted: place errors (e.g., *k* for *t*), voicing errors (e.g., *v* for *f*), and affrication errors (e.g., *jr* for *dr*). Errors on any of the 25 words could include a feature error.

These specific types of errors were selected for analysis because previous research has suggested that these errors are commonly made by beginning spellers (e.g., Treiman, 1993). However, such errors have not been systematically documented at higher grade levels. We predicted that, if there were differences in spelling levels as a function of instructional method, the less skilled group would be more likely to commit these types of errors. We note, however, that the stimuli were not constructed specifically to examine these types of errors, and as a result there may not be sufficient instances to each category to measure children's spelling skills in each of these areas reliably. Nonetheless, this post-hoc analysis may be instructive in terms of the trends and implications for future research.

A similar coding scheme was used for the nonwords. Nonword spellings were classified into one of the three categories: constrained, unconstrained, or phonologically inaccurate. The preservation of skeletal structure and the orthographic acceptability of each nonword spelling was also coded. Finally, errors involving postvocalic nasal clusters, other consonant clusters, and the phonological features of place, voicing, and affrication were also examined.

A total of 100 unique misspellings of words and 100 unique spellings of nonwords was rescored by a second rater. The items (types) were selected randomly with the constraint that there be at least two different instances of misspellings for each of the word and nonword stimuli. These items (types) represented 32% of all misspellings (tokens) and 30% of all nonword spellings (tokens). Agreement between the two raters across the categories (constrained, unconstrained, orthographically acceptable, and preservation of phonological skeleton structure) ranged from 100 to 96%. Discrepancies were resolved by discussion.

RESULTS

The two groups were well matched on two important dimensions. First, the educational levels of their mothers and fathers were similar (all chi-square analyses were nonsignificant; see Table 1), reflecting similar socioeconomic backgrounds. Thus, although the phonics group attended a private (parochial) school

Table 1. *Background characteristics of whole language and phonics children*

	Whole language ($n = 54$)	Phonics ($n = 22$)
Age (months)	110 (4)	105 (2)
Mother's educational background		
High school or less (%)	51	62
College or more (%)	49	38
Father's educational background		
High school or less (%)	33	47
College or more (%)	67	53
Word recognition raw score	61 (9)	63 (8)
Word recognition percentile	47 (22)	60 (22)
< 25%	27	5
26%–50%	27	32
51%–75%	36	22
> 76%	9	22

Note: Standard deviations are in parentheses.

whereas the whole language children attended public schools, the phonics group was not more advantaged. Second, as shown in Table 1, the two groups of children had similar raw scores on the word recognition subtest of the Woodcock Reading Mastery Test–Revised, $F(1, 74) = 1.22$, $p < .27$.

Although the raw scores on the word recognition subtest were similar, the whole language children's percentile scores on the word recognition subtest were significantly lower than those of the phonics children. On average, the phonics group scored at the 60th percentile, whereas the whole language group scored at the 47th percentile, $F(1, 74) = 4.24$, $p < .05$. The difference between the raw score and percentile score results is due to the fact that, at the time of testing, the whole language children had three months more schooling and were approximately five months older than the phonics children. As shown in Table 1, the between-group difference in percentile scores reflects the fact that, as compared with the phonics group, there were relatively more whole language children who scored below the 25th percentile and fewer who scored above the 76th percentile.

On the basis of previous literature documenting the high association between reading and spelling skills (see Bruck & Waters, 1990, for a review), one would predict that, because the phonics and the whole language children had the same level of word recognition skills (in terms of raw scores), they would also have comparable levels of spelling skills.

In order to compare the spellings of the whole language and phonics groups, separate one-way analyses of variance were carried out on the following dependent variables: correctly spelled words, percentage of misspellings that preserved the skeletal structure of the target word, and percentage of misspellings that were orthographically acceptable. A two-way analysis of variance with re-

Table 2. *Spelling of words*

	Whole language ($n = 54$)	Phonics ($n = 22$)
Correct spellings (%)	55 (21)	69 (17)
Phonologically accurate errors		
Constrained errors (% of errors)	39 (29)	42 (34)
Unconstrained errors (% of errors)	18 (14)	14 (12)
Misspellings preserving skeletal structure (% of errors)	69 (22)	66 (29)
Orthographically acceptable errors (% of errors)	71 (20)	72 (28)
Postvocalic nasal cluster errors (%)[a]	7 (14)	10 (17)
Other consonant cluster errors (%)[b]	10 (14)	8 (10)
Consonant feature errors (%)[c]	2 (3)	1 (2)

Note: Standard deviations are in parentheses.
[a]The percentage score was derived by dividing the number of preconsonantal nasal cluster errors by the number of words containing nasal consonant clusters.
[b]The percentage score was derived by dividing the number of other consonant cluster errors by the number of words containing "other" consonant clusters.
[c]The percentage score was derived by dividing the total number of feature errors by the number of words.

peated measures was carried out on the measures of phonological accuracy. The independent variables were instructional group and type of measure (constrained vs. unconstrained). The data are summarized in Table 2.

The phonics group spelled more words correctly than the whole language group, $F(1, 74) = 6.72$, $p < .01$. However, with one exception, the groups did not differ significantly in the types of spelling errors made. That is, both groups made the same proportions of phonologically accurate errors (both constrained and unconstrained), orthographically acceptable errors, and errors preserving the phonological skeleton of the target word. In addition the two groups did not differ in terms of errors on postvocalic nasals in clusters and errors on other consonant clusters. The one exception was a significant between-group difference on the number of feature errors made, with the whole language group making more feature errors on consonants than their phonics peers, $F(1, 74) = 4.00$, $p < .05$. However, the rate of feature errors was low for both groups, with most children not making such errors. Furthermore, most of the feature errors occurred on the word *dramatic* where the most common errors involved affrication, as in "gramadic" and "jermadik." In contrast to the low rate of feature errors, consonant cluster errors (on nasals and other clusters) did occur with some frequency in this grade 3 sample. Thus, children misspelled "dive" for *drive*, "tansparent" for *transparent*, and "tet" for *tent*.

Although phonological accuracy of misspellings is a common measure in spelling studies, the manner in which phonological accuracy is measured varies across studies, with some researchers using a constrained method and others using an unconstrained method. It is not known which of these methods best correlates with spelling proficiency. Therefore, in order to address this issue,

Table 3. *Spelling of nonwords*

	Whole language ($n = 54$)	Phonics ($n = 22$)
Phonologically accurate spellings		
Constrained (% of all items)	48 (19)	57 (18)
Unconstrained (% of all items)	14 (7)	10 (6)
Spellings preserving skeletal structure (% of items)	77 (16)	81 (16)
Orthographically acceptable spellings (% of items)	79 (12)	82 (11)
Postvocalic nasal cluster errors (%)[a]	16 (23)	18 (27)
Other consonant cluster errors (%)[b]	12 (14)	14 (10)
Consonant feature errors (%)[c]	4 (4)	4 (4)

Note: Standard deviations are in parentheses.
[a]The percentage score was derived by dividing the number of preconsonantal nasal cluster errors by the number of nonwords containing nasal consonant clusters
[b]The percentage score was derived by dividing the number of other consonant cluster errors by the number of nonwords containing "other" consonant clusters
[c]The percentage score was derived by dividing the total number of feature errors by the number of nonwords.

correlations were computed between the number of correctly spelled words and the constrained and unconstrained scores. High rates of constrained errors were positively correlated with high numbers of correct spellings ($r = .70$, $p < .01$). However, high rates of unconstrained errors were associated with low numbers of correct spellings ($r = -.34$, $p < .01$). Thus, it appears that, for grade 3 children, constrained errors were associated with competent spelling, whereas unconstrained errors were associated with less competent spelling.

Table 3 presents the performance of the whole language and phonics group on the various measures of nonword spelling. A two-way analysis of variance with repeated measures was carried out on the phonologically accurate spellings.[1] The independent variables were instructional group (whole language vs. phonics) and scoring method (constrained vs. unconstrained), with repeated measures on the last factor. There was a main effect of scoring method, $F(1, 74) = 243.11$, $p < .001$, and a significant interaction between instructional group and scoring method, $F(1, 74) = 6.46$, $p < .01$. Post-hoc tests revealed that the phonics group produced more constrained spellings than the whole language group, but the whole language group produced more unconstrained spellings than the phonics group. Thus, the phonics group produced more spellings with conventional phoneme–grapheme correspondences.

One-way analyses of variance were carried out on the remaining measures. As in the analyses of the word data, the two groups did not differ on the measures of preservation of skeletal structure and orthographic acceptability. Also, as Table 3 shows, the two groups did not differ in the number of errors on postvocalic nasal clusters, other clusters, or consonant features. As was found in the parallel analyses of the word data, a significant number of such errors were made at the grade 3 level. Children sometimes misspelled *trive* as "tive"

or "tirv" or *nent* as "nat." Again, affrication was a common error, as when *drayer* was misspelled as "graer" or "jraer."

In order to determine which measure of nonword phonological accuracy was better associated with spelling proficiency, constrained and unconstrained nonword spellings were each correlated with the number of correctly spelled words. The pattern of results was the same as for the word data. Children with high rates of constrained nonword spellings also had high accuracy rates on the word spelling test ($r = .78$, $p < .01$). However, children with high rates of unconstrained nonword spellings had low accuracy rates on the word spelling test ($r = -.24$, $p < .05$).

DISCUSSION

There are several important findings, the first of which pertains to word recognition skill. Although the children in the whole language and phonics groups were matched on raw word recognition scores, the whole language group had lower percentile scores because they were somewhat older than the children in the phonics group. It was not simply that the whole language children scored at lower but acceptable levels for their age. Rather, the average word recognition score was below the 50th percentile, and a relatively large proportion of the children scored below the 25th percentile (a criterion sometimes used for classifying children as "very poor" readers). Given the relatively high socioeconomic status of these children, these results are a cause for concern.

The second important finding is that the children in the whole language program group did not spell appropriately for their age level or their reading level. Despite comparable levels of raw word recognition scores, the whole language children's word and nonword spelling scores were significantly lower than those of the phonics children. These data suggest that the absence of explicit phonics instruction has a greater negative impact on spelling than on word recognition. If the impact on spelling and word recognition were similar, the whole language children should have performed similarly to the phonics children on the spelling measures, just as they did on the word recognition measure.

Third, the data indicate that the whole language and phonics groups did not differ in their attempt to use phoneme–grapheme information when spelling. The groups were similar in terms of their ability to preserve the skeletal structures of spoken words and nonwords in spelling. This similarity reflects their ability to segment the spoken items and their attempt to assign graphemes to phonemes (i.e., use phoneme–grapheme information). Both groups produced a similar number of phonologically motivated spellings for the nonwords (i.e., the total number of phonologically accurate constrained and unconstrained spellings was similar for the two groups), and there were no differences in the percentage of word errors that were phonologically motivated (by either the constrained or unconstrained system). Again, these findings suggest that whole language and phonics children attempted to use phoneme–grapheme information to the same extent. It appears that both groups were equally knowledgeable about the basics

of phoneme–grapheme correspondences: they knew which letters represent consonants and which letters represent vowels.

Fourth, the phonics and whole language children's spellings did not differ in orthographic acceptability. Although some principles of orthographic acceptability may be taught in phonics programs, perhaps exposure to print is sufficient for children to internalize many of these principles. Supporting this view, Treiman (1993) found that grade 1 children in a whole language program picked up certain orthographic patterns on their own, without explicit teaching. For example, these children appeared to know that *ck* may not appear at the beginning of an English word, as they rarely produced misspellings such as "ckut" for *cut*.

Finally, despite the similarities in some of the spelling strategies, the phonics and whole language groups differed in their knowledge of conventional phoneme–grapheme correspondences – specifically, knowledge about the contexts in which various correspondences are used. Thus, on the nonword spelling test, the whole language children produced fewer positionally constrained spellings than the phonics children. This difference is not a minor point, for it accounts for the poorer spelling of the whole language group. Positional constraints must be a fundamental component of the English spelling system, or else George Bernard Shaw's legendary "ghoti" would be a common misspelling for *fish*. This gap in the whole language children's knowledge probably reflects the lack of systematic instruction on the basic spelling patterns of English.

Of course, there are certain limitations to the present study – limitations which can be easily addressed by future research. First, it is important to replicate the results with children from different whole language and phonics classrooms. Although the children in this study were drawn from a number of whole language schools and classrooms, the phonics children came from a single grade 3 classroom. These children had, however, attended different classrooms prior to this grade. In addition, the sample size for the phonics group was relatively small. Second, it would be important in future studies to construct spelling lists so that they intensively sample specific spelling patterns that are potentially important markers of spelling ability. Thus, although we examined spellings of clusters, the stimuli were not strictly equated on a number of important dimensions (e.g., initial vs. final clusters). Third, our data may overestimate the skills of whole language children. On the one hand, we suspect that had we selected age-matched groups (rather than reading-matched groups), the whole language children would have performed much more poorly than they did in the present study. It is also the case that, if we had examined children's spellings in the contexts of their daily writing assignments (rather than examining errors in dictated lists of words), the whole language children's spellings would have been much worse relative to their phonics instructed peers. When children's attention is focused on meaning and creativity and spelling is not an automatized skill, more errors are likely to occur.

Despite these limitations, our results suggest some important methodological and substantive conclusions. One concerns the methods used to score phonologically accurate spellings and misspellings. In the past, researchers have often interchangeably used constrained and unconstrained systems, with the two

sometimes producing different results (e.g., Bruck & Waters, 1988). To our knowledge, there has been little discussion of which system should be used in which situation. The choice of system often seems arbitrary. The present results indicate that, for 8-year-old children who have developed a number of skills, the constrained system most accurately measures "competence," whereas the unconstrained system reflects the use of less mature strategies. Supporting this view, constrained spellings on both the word and nonword spelling tasks were associated with good spelling of real words. The opposite pattern was obtained for the unconstrained spelling measures on both the word and nonword spelling tests. These were associated with poor ability to spell real words.

A second implication is that, even in grade 3, children make a number of errors that were thought to be characteristic of very beginning spellers. Thus, although the error rates were not high, children in both the whole language and phonics programs did make errors on postvocalic nasals in final clusters and on other types of syllable-final and -initial clusters. These errors mainly involved the omission of one of the consonants in the cluster. Such errors have previously been documented in preschool inventive spellers (Read, 1975) and in kindergarten and grade 1 children (e.g, Treiman, 1991, 1993; Treiman, Zukowski, & Richmond-Welty, 1995). Moreover, certain feature errors similar to those previously documented in younger children (Read, 1975; Treiman, 1985) were noticed in the present study. For example, 31% of the whole language children and 18% of the phonics children affricated the /dr/ in *dramatic* to produce such spellings as "grematic" or "jermatic." There was only one such error for the monosyllabic word *draw* – "graw." The higher error rate on the multisyllabic compared with the monosyllabic word is consistent with work by Snowling (1994). In the future, therefore, it would be important to include both long and short words when assessing children's spelling abilities.

To conclude, our data suggest that, as compared with children in phonics programs, children in whole language programs do not make the same progress on basic skills of word recognition and, most important for the present study, spelling. It seems that these deficits could be remediated without challenging fundamental aspects of the whole language program, such as writing in personal journals and reading interesting materials. Daily lessons in which children are taught systematic relationships between sounds and spellings for both word recognition and spelling could easily right the balance. Previous research has suggested that this add-on component increases the reading skills of whole language children (Castle et al., 1994; Iverson & Tunmer, 1992; Rubin & Eberhardt, 1996). Our results suggest that it would have a similar effect on spelling.

APPENDIX

WORD AND NONWORD STIMULI

Words	Nonwords
hard	bard
tent	nent
big	hig
drive	trive
song	jong
hope	vope
shop	shap
feet	reet
voice	soice
field	bield
edge	quidge
draw	plaw
player	drayer
pencil	hencil
round	lound
afraid	aglaid
paper	maper
magic	fagic
receive	beceive
finally	cheenally
cleverness	blemerness
surprising	turclising
protective	frojective
transparent	planscarent
dramatic	bracatic

ACKNOWLEDGMENTS

This research was supported by Grants 0GP000A1181 from the Natural Sciences and Engineering Research Council to M. Bruck, 410-91-1936 from the Social Sciences and Humanities Research Council to F. Genesee, 94-ER-2015 from Fonds pour la Formation de Chercheurs et l'Aide à la Recherche to M. Bruck and F. Genesse, and SBR-9408456 from the National Science Foundation to R. Treiman.

NOTE

1. The constrained and unconstrained nonword measures are not exactly comparable to the constrained and unconstrained word measures. For the nonwords, the constrained spellings might be regarded as the most "correct" spellings and thus more comparable to correct word spellings than to constrained word misspellings. However, the percentage of constrained nonword spellings cannot be directly compared with the percentage of correct word spellings because there are a number of possible spellings that are reasonable for the nonwords, whereas each word has only one possible spelling. For these reasons, the word and nonword data were not directly compared in statistical analyses.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Bruck M., & Waters, G. (1988). Analysis of the spelling errors of children with discrepancies between their reading and spelling skills. *Applied Psycholinguistics, 9,* 77–92.

(1990). An analysis of the component spelling skills of good readers–poor spellers. *Applied Psycholinguistics, 11,* 425–437.

Butler, S. (1994). The writing connection. In V. Froese (Ed.), *Whole language: Practice and theory* (2nd ed., pp. 149–198). Scarborough, Ontario: Allyn & Bacon.

Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book.* Boston: Houghton Mifflin.

Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program. *Journal of Educational Psychology, 86,* 350–359.

Edwards, J. (1985). Spelling corrections alter children's voices. *Highway One: Canadian Journal of Language Education, 8,* 6–14.

Eldredge, J. L., & Baird, J. (1996). Phonemic awareness training works better than whole language instruction for teaching first graders how to write. *Reading Research and Instruction, 35,* 193–208.

Graham, S., & Harris, K. (1994). The effect of whole language on children's writing: A review of literature. *Educational Psychologist, 29,* 187–192.

Iverson, S., & Tunmer, W. E. (1992). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology, 85,* 112–126.

Read, C. (1975). *Children's categorization of speech sounds in English research* (Report No. 17). Urbana, IL: National Council of Teachers of English.

Rubin, H., & Eberhardt, N. (1996). Facilitating invented spelling through language analysis instruction: An integrated model. *Reading and Writing: An Interdisciplinary Journal, 8,* 27–43.

Snowling, M. J. (1994). Towards a model of spelling acquisition: The development of some component skills. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 111–128). Chichester: Wiley.

Stahl, S. A., McKenna, M. C., & Pagnucco, J. R. (1994). The effects of whole language instruction: An update and a reappraisal. *Educational Psychologist, 29,* 175–185.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59,* 87–116.

Treiman, R. (1985). Phonemic awareness and spelling: Children's judgments do not always agree with adults'. *Journal of Experimental Child Psychology, 39,* 182–201.

(1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology, 83,* 346–360.

(1993). *Beginning to spell: A study of first-grade children.* New York: Oxford University Press.

Treiman, R, Zukowski, A., & Richmond-Welty, E. D. (1995). What happened to the "n" of sink? Children's spellings of final consonant clusters. *Cognition, 55,* 1–38.

Uhry, J., & Shepherd, M. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects of several measures of reading. *Reading Research Quarterly, 28,* 218–233.

Varble, M. (1990). Analysis of writing samples of students taught by teachers using whole language and traditional approaches. *Journal of Educational Research, 83,* 245–251.

Venezky, R. (1980). *The structure of English orthography.* The Hague: Mouton.

Wilde, S. (1989). Looking at invented spelling: A kidwatcher's guide to spelling. Part I. In K. Goodman, Y. Goodman, & W. Hood (Eds.), *The whole language evaluation book* (pp. 213–226). Portsmouth, NH: Heinemann.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Test–Revised.* Circle Pines, MN: American Guidance Service.

Zarry, L. (1991). *Literacy through whole language.* Winnipeg: Peguis Publishers.