Learning Disabilities Research & Practice, *26*(1), 15–27

An Examination of a Small-Group Decoding Intervention for Struggling Readers: Comparing Accuracy and Automaticity Criteria

Roxanne F. Hudson, Carole Isakson, and Taylor Richman
University of Washington

Holly B. Lane and Stephanie Arriaza-Allen
University of Florida

In this study, we compared methods to improve the decoding and reading fluency of struggling readers. Second-grade poor readers were randomly assigned to one of the two practice conditions within a repeated reading intervention. Both interventions were in small groups, were 20–28 min long, took place 2–4 days per week, and consisted of phonemic awareness training, letter sound practice, and practice in word families. Students in the accuracy condition ($n = 27$) practiced each page until they reached 98 percent accuracy while students in the accuracy + automaticity condition ($n = 29$) practiced until they reached rate (30–90 cwpm) and accuracy criteria. Hierarchical linear modeling revealed no differences between practice conditions in decoding accuracy, reading comprehension, and grade-level text reading fluency. Significant differences favoring the accuracy + automaticity group were found in measures of decoding automaticity.

Early intervention for reading problems is generally considered critical for later student success. There is strong evidence that improving children's phonemic awareness, knowledge of letter–sound correspondences, and decoding skills will improve their reading (e.g., National Reading Panel, 2000; Torgesen & Hudson, 2006). Students with significant reading difficulties tend to struggle with decoding and word-level reading, greatly impairing their ability to comprehend what they read (Chard, Vaughn, & Tyler, 2002). Decoding accuracy is necessary but not sufficient for proficient reading; readers need to develop decoding skills to a level of automaticity. Automaticity is a necessary component of skill, and higher-level aspects of skill cannot be acquired until lower-level aspects have become automatized (Bryan & Harter, 1899; Laberge & Samuels, 1974). If reading subskills are performed automatically, then comprehension and other higher-order aspects of reading can function effectively at the same time (Samuels & Flor, 1997).

The development of effective intervention methods for improving reading subskills is an ongoing challenge. Previous research has examined several aspects of such instruction, including (1) comparing practice with words in isolation to practice with words in context, (2) the importance of familiarity with within-word patterns, (3) the role of goal setting and corrective feedback during reading instruction, (4) the relative impact on the type of cues given during reading practice, and (5) the effects of various performance criteria on retention of reading-related skills.

Requests for reprints should be sent to Roxanne Hudson, Ph.D., Area of Special Education, University of Washington, Box 353600, Seattle, WA 98195. Electronic inquiries should be sent to rhudson@u.washington.edu.

Results from studies examining the role of practice with isolated words compared to practice with words in context have been mixed. In some studies, direct comparisons of both methods have shown no significant difference (LeVasseur, Macaruso, & Shankweiler, 2008; Levy, 2001). In other studies, practice of isolated words has produced improvements in reading rate and accuracy (Fleisher, Jenkins, & Pany, 1979; Levy, Abello, & Lysynchuk, 1997; Tan & Nicholson, 1997), while others have found that repeated practice of connected text is superior to word lists (Dahl, 1979; Daly & Martens, 1994; Martin-Chang & Levy, 2005; Morgan & Sideridis, 2006). It is important to closely examine how these approaches to instruction relate to the specific needs of students, because different students may struggle with different levels of reading, requiring different areas of focus during interventions.

Automaticity in recognition of phonograms (i.e., letter groups within a word that share a pattern across words) is a feature of the more advanced word recognition (Ehri, 1992). Without knowledge of patterns across words, readers are not able to move to more advanced, efficient decoding (Ehri, 2002). For example, a child who decodes using individual letters would take much longer to read a word like *uncomfortable* than one who can decode using larger chunks such as *un-com-fort-able*. Children read words with frequent or familiar rimes (a vowel plus syllable ending) more accurately than those with infrequent or unfamiliar ones (McKay & Thompson, 2009), and instruction in how to use rimes as analogies to assist with word recognition has been successful with young children (Goswami & Mead, 1992; Levy & Lysynchuk, 1997; Walton & Walton, 2002). In addition, English is more regular at the level of rimes and larger chunks than at the phoneme–grapheme level (Kessler & Treiman, 2003; Moats, 2000), making sound–symbol relationships at that level more predictable and useful in reading words. The rime -ake is almost always pronounced with a long a sound while the letter a can spell several different sounds. Readers need to develop context-sensitive mappings of relationships between phonemes and graphemes as well as larger units such as rimes, morphographs, and affixes, to become fluent decoders and readers (Brown & Deavers, 1999; Berninger, Abbott, Vermeulen, & Fulton, 2006).

An important aspect of intervention designed to increase reading fluency is goal setting and corrective feedback. In a meta-analysis of single-subject studies of reading fluency, Morgan and Sideridis (2006) found that, among instructional methods, programs that focused on goal setting (goal setting plus feedback, and goal setting plus feedback and reinforcement) were most effective at improving struggling readers' rate and accuracy. They suggest that these programs provide students with a set plan and error correction, which supports self-regulation and motivation. Therrien (2004) also found that students who received corrective feedback from an adult during reading or immediately after, substantially improved in reading rate and accuracy (average effect size = 1.37).

Another element worth considering is the type of cues given during reading tasks, because the cue provided can affect the student's focus and performance. For example, O'Shea, Sindelar, and O'Shea (1987) compared the use of fluency cues versus comprehension cues given to students with and without learning disabilities. O'Shea et al. found that students with learning disabilities failed to adapt their reading style to the cues presented to them. Those who were asked to read for comprehension were able to remember more, but continued to read as fast as those who were cued for speed. In contrast, students without learning disabilities had the ability to modify their reading style according to the cues presented. Other researchers have found different results. For example, Therrien (2004) reported the highest effect size (.94 for fluency and .67 for comprehension) when students were cued to read for both speed and comprehension. Pressley, Hilden, and Shankland (2006) found no difference in students' reading fluency or the correlation between fluency and comprehension based on the cue to "do your best reading," "read as quickly as possible," and "it is important you read to understand the story." Hudson, Torgesen, and Schatschneider (2006) found that second-graders read faster when asked to read "as fast as you can" than when asked to read "as carefully as you can," but the correlations with a standardized comprehension test were not significantly different from each other (.61 with speed condition, .58 with accuracy condition).

Haughton (1972) was among the first to determine that accuracy in a skill was not equivalent to mastery and that automaticity was essential for retention and success in more difficult skill levels. Binder, Haughton, and Bateman (2002) lament that, too often, mastery to a given level of accuracy with a skill is the only goal of instruction. They suggest that automaticity is at least as important "because it allows students to progress smoothly through the learning process, building each successive layer on a previous layer of fluent prerequisite skills and knowledge" (p. 5). Although we found no direct comparisons of accuracy and automaticity performance criteria in decoding skills, at least one study has compared the effects of accuracy and automaticity criteria on the application and retention of a composite skill. In Bucklin, Dickinson, and Brethower's (2000) study, college students were taught associations between Hebrew symbols and nonsense syllables and between nonsense syllables and Arabic numerals. They found that training with an emphasis on automaticity resulted in higher response rates and better retention of both the component and composite skills.

Differences in comprehension and reading fluency between good and poor readers can be at least partially attributed to differences in decoding automaticity (Perfetti & Hogaboam, 1975; Samuels & Flor, 1997; Torgesen, 1986; Torgesen, Rashotte, & Alexander, 2001), and decoding automaticity depends on automatic processing at the sublexical and lexical levels (Hudson, Pullen, Lane, & Torgesen, 2009). Automaticity at the letter, letter pattern, and word levels is critical because reading processes share limited-capacity processing resources in working memory, and a lack of automaticity at the lower levels restricts full processing at the higher ones (Perfetti, 1985). Little evidence exists, however, to indicate whether practice to develop these necessary skills is effective. Given the benefits in reading rate and accuracy in unpracticed texts from repeated reading practice of isolated words (Fleisher et al., 1979; Levy et al., 1997; Tan & Nicholson, 1997), and given that decoding accuracy is the focus of early reading instruction (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998), we determined that investigation of the effects of repeated practice reading word parts was warranted. Furthermore, given the conflicting findings related to the types of cues students receive, we determined that research was also needed to compare intervention with cues focused on developing accuracy with intervention with cues focused on both accuracy and automaticity.

RESEARCH QUESTIONS

We conducted an experimental study of small-group instruction in phonemic awareness, letter–sound correspondences, and word families with struggling second-grade readers. Two conditions were compared: repeated reading of a page of letter sounds and a page of words until an accuracy criterion (98 percent) was reached or repeated reading of the same materials until a rate and accuracy criterion was met (30–90 cwpm). The main purpose of this study was to test whether practice in lower-level skills, in the absence of text-level practice, would transfer to mid-level skills such as decoding and the upper-level skills of oral reading fluency and reading comprehension. Specific questions we addressed were (1) Is either practice condition (reading to an accuracy or to an accuracy plus rate criterion) more effective than the other for improving decoding accuracy, decoding automaticity, or reading comprehension after controlling for curriculum mastery, or how many pages of words were mastered? (2) After controlling for curriculum mastery, what is the effect of practice condition on growth in reading rate and accuracy in grade-level text? (3) Are there differential responses to the practice conditions due to disability status?

METHOD

Participants

Seven schools in three school districts in Washington and Florida participated in the study. Second-grade teachers in participating schools were asked to nominate readers who struggled in text reading rate and accuracy. Those students whose parents gave informed consent were screened using three second-grade passages from the Oral Reading Fluency subtest of the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS ORF; Good & Kaminski, 2002) and the Picture Vocabulary subtest of the *Woodcock Johnson Test of Academic Achievement, III* (WJ III; Woodcock, McGrew, & Mather, 2001). Students whose median score on the DIBELS ORF was at or below the 35th percentile and at or above 45th percentile on the vocabulary measure were invited to participate in the study. These criteria were established to ensure the participants had a problem with reading fluency and would have a sufficient command of English vocabulary to benefit from reading text in English.

The parents of 58 students gave consent for participation (55 percent boys). Demographics obtained from school records for each intervention group are reported in Table 1. Sixty-nine percent had no identified disability label while one child's disability status was unknown. Fourteen percent were eligible for services in English as a Second Language (ESL) in Washington State, which meant they scored at or below Level 3 on the Washington Language Proficiency Test II, which is given to every student whose family self-identifies as speaking a language other than English at home. This test assesses listening, speaking, reading, and writing in English with an emphasis on written communication. According to the Washington Office of Superintendent of Public Instruction (n.d.), a student who scores at Level 3, "Reads both long and short connected texts with understanding. Writes simple essays with standard conventions, organization and detail. Uses figurative and idiomatic language in discussions of academic content and ideas." No students were eligible for ESL services in Florida. Forty-three percent were identified in school records as African American, 24 percent as Caucasian, 8.6 percent as White Hispanic, 6.9 percent as Indigenous Hispanic, 6.9 percent as Asian, 5.2 percent as Vietnamese, 1.7 percent as African, and 1.7 percent as Korean. One student's ethnicity is unknown. Two students moved during the study; complete data for 56 students are reported in this article.

TABLE 1
Participant Demographics by Group

	Accuracy	*Accuracy + Automaticity*	*Total*
Race/ethnicity			
White	7 (26%)	7 (23%)	14 (25%)
African American	11 (41%)	14 (47%)	25 (44%)
African	0	1 (3%)	1 (1.7%)
Hispanic	4 (15%)	5 (17%)	9 (16%)
Asian	5 (19%)	3 (10%)	8 (14%)
Female	9 (3%)	16 (53%)	25 (44%)
Special education			
Emotional/behavioral disability	0	1 (3%)	1 (1.7%)
Specific learning disability	6 (22%)	5 (17%)	11 (19%)
Multiple disability	0	2 (7%)	2 (4%)
Other disabilities	1 (4%)	2 (7%)	3 (5%)
English language learners	4 (15%)	4 (13%)	8 (14%)
Average number of sessions	34	34	34

Intervention Procedure

The available literature suggests that repeated reading intervention takes place at least three times a week for 10–20 min each session (Chard et al., 2002). In each session, students should work with an adult who models fluent reading (Therrien, 2004) and provides corrective feedback (Chard et al., 2002). It is suggested that students be asked to read as many times as necessary until they meet a set criterion, while implementing some type of motivational component, including goal setting and reinforcement (Morgan & Sideridis, 2006). The focus of the intervention should be based on the specific needs of the students, particularly when working with students with significant reading problems.

Participants who met the criteria were randomly assigned within school to either an accuracy condition ($n = 27$) or an automaticity (accuracy and rate) condition ($n = 29$). Children who were randomly assigned to condition were then placed in small groups based on school schedules and both interventions occurred in groups of two to four students, were 20–28 min long, and took place two to four times a week. All but two tutors worked in both conditions and all tutors taught multiple small groups. The goal for each condition was 40 sessions, however, due to school schedules, this was not always met. Children in both conditions had a mean number of sessions of 34. In the accuracy condition, the range was 17–40 ($SD = 5$) while children in the accuracy + automaticity condition had a range of 27–40 ($SD = 4$). The goal was for the intervention to serve as a supplement to ongoing reading instruction. Thus, the majority of tutoring sessions took place during other instructional times such as physical education, science, or art. No student missed special education or ESL services to participate in the study. Each condition had three steps: a phonemic awareness warm-up, isolated letter sound practice, and word family practice.

Tutors

The tutors varied in experience and educational preparation. All had bachelor's degrees and experience teaching or tutoring children. One was a certified teacher; three were graduate students in special education, school psychology, and law; three worked as Amercorps volunteers in the school where they tutored; and three were parents who had experience with other tutoring programs.

Intervention Materials

The two conditions used the same materials. All participants practiced the same words in the phonemic awareness warm-up and used the same pages of isolated letter sounds and word families. The phonemic awareness materials came from Blachman, Ball, Black, and Tangel (2000) and consisted of plastic disks that students manipulated and a paper work space for each one. There were 16 isolated letter pages that began with continuous sounds and short vowels, moved to include stop sounds and consonant digraphs, and ended with vowel digraphs and uncommon consonant digraphs (see Appendix A for progression). The first page had five sounds repeated to make 35 on the page (e.g., a, s, n, r, m) and each subsequent page increased by five sounds to the ultimate goal of 60 per page. Letter names were not taught because of the advanced grade level of the participants and the project's focus on the alphabetic principle. There were 40 word pages that followed the same progression of sounds as the isolated sound pages in the form of three to four word families per page. The first page began with words repeated to make 25 on the page (e.g., ram, at, man, fan, am) and increased by five words a page to the ultimate goal of 100 words per page (see Appendix A for progression). While the word families repeated across pages, the individual words varied from practice page to practice page. For example, on page 7, the -ack family is introduced with tack, back, pack, and shack. On page 8, -ack is repeated using the words sack and rack as exemplars.

Accuracy + Automaticity Condition

In this condition, the goal was to develop automaticity with letter sounds and decodable words organized into word families. Students were cued to read both accurately and quickly. In Step 1, the phonemic awareness warm-up, students engaged in "Say it and Move It" (Blachman et al., 2000) as a whole group. With modeling and corrective feedback, they segmented and blended two-phoneme words and gradually increased in difficulty to segmenting and blending words with five phonemes.

In Step 2, the participants individually practiced a page of isolated letter sounds with modeling from the tutor for the new sounds and corrections as needed. With their tutor, students inspected their graphs and set a rate and accuracy goal that they would accomplish. Next they read the sounds aloud to the intervention tutor for 1 min while the tutor kept track of errors. Then the tutor gave corrective feedback on both rate and accuracy, calculated the number of correct sounds read in the minute, and graphed the score with the student. If the student read to the end of the page in 1 min with two or fewer errors, practice began on the next page during the next session. If not, the student repeated the page the next session and received feedback to read more quickly and/or more accurately. While the tutor worked with each individual, the other students practiced quietly to themselves. The accuracy criterion was 2 or fewer errors per page and the rate criterion increased in a stairstep fashion from 35 correct sounds per minute (cspm) to 60 cspm. The intervention began with a very low rate criterion based on the authors' previous experience with struggling readers who are frequently very slow. Beginning with a goal at the top of the criterion would have been daunting to a student who said letter sounds at 30 cpm.

In Step 3, the activities of Step 2 were repeated with a word page. The same accuracy criterion of 2 or fewer errors per page was used and the rate criterion increased from 25 to 100 cwpm. This final criterion level was chosen based on a slightly higher than 50th percentile on the Hasbrouck and Tindal (2006) spring norms for second grade. At the end of the session, students who reached their goal added a small star sticker to a larger group star. When the group star was filled, the entire group received a large shiny paper star with their name on it. Motivation was provided through the use of goal setting, graphing progress, and earning stars.

Accuracy Condition

In the accuracy condition, participants were cued to read correctly regardless of speed and the same three steps were followed with a goal of developing accuracy. Step 1 was identical to the accuracy + automaticity condition.

In Step 2, the participants individually practiced a page of isolated letter sounds with modeling from the tutor for new sounds and correction as needed, looked at the graph with the intervention tutor, and set an accuracy goal that they would accomplish, and then read the sounds aloud to the tutor for 1 min while the tutor kept track of errors. Then the tutor gave corrective feedback on accuracy, calculated the percentage of words read correctly, and graphed the score with the student. If the student met the accuracy criterion (95 percent for the first page, then increasing to the ultimate goal of 98 percent by the third page), then the student moved to the next page. For example, if a student read 10 sounds, but all were correct, the student moved to the next page, regardless of the fact he or she did not reach the end of the page. If the student did not reach the accuracy goal, he or she repeated the page the next time. The accuracy criterion of 98 percent was established because it was necessary for the students to master the sounds, but the authors wanted to leave a little for the occasional error.

In Step 3, the activities of Step 2 were repeated with a word page. Again, the cues focused on reading correctly, not quickly. As with the other condition, each student who made his or her goal added a small star sticker to a larger group star at the end of the session. When the group star was filled, the entire group received a large shiny paper star with their name on it. Motivation was provided through the use of goal setting, graphing progress, and earning stars.

Implementation Fidelity

The first author trained all tutors in a 2-h session to introduce the materials, timers, and procedures. The first author modeled each step of the automaticity procedure and all tutors practiced each of the steps. All tutors practiced the procedures independently and then attended a second 2-h training session facilitated by the first author in Washington and the second and third authors in Florida. In this second session, tutors asked implementation questions and the accuracy condition was introduced and practiced. Finally, before any tutor could work with children, he or she tutored the first or third author and earned a score of 90 percent or better on a treatment fidelity checklist. During the second or third session with children, each tutor was observed and given feedback using the fidelity checklist by the first or third authors. Correction consisted of graphing errors and occasional reference to number of words read by two tutors in the accuracy condition. This was corrected and did not occur after this initial observation. Each tutor was observed at least two more times using the checklist. Fidelity was consistently high; we found no failure to implement the assigned treatment within the small groups. In addition, weekly meetings were held to address implementation difficulties and answer tutor questions. Three of the tutors needed additional coaching and training on managing student behavior.

Measures

All participants were individually assessed just before they began and just after they finished the intervention. All raw data protocols were scored a second time by a research assistant to ensure correct scoring and all assessment data were entered twice by different research assistants and discrepancies corrected on a case-by-case basis.

Decoding Accuracy

Decoding accuracy was measured by the decoding subtest of the *Kaufman Test of Educational Achievement II* (Kaufman & Kaufman, 2004). Students were asked to read increasingly difficult decodable nonsense words (e.g., ip, rell, trowful, dompest) until they missed four in a row. Standard scores were determined using grade-level norms.

Decoding Automaticity

Two assessments were used to measure decoding rate and accuracy. In the phonemic decoding efficiency (PDE) subtest of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), participants were asked to read as many nonsense words presented in list format as possible in 45 s. Both form A and form B were given and the average of the two scores used in the analysis; obtained alternate form reliability for the pretest was .86 and for the posttest, .90. Standard scores were determined using grade-level norms.

For the second assessment of Complex Nonsense Word Decoding, we constructed two forms of 64 randomly ordered monosyllabic nonsense words (e.g., bost, stribe, jipe, spurge, cheel) presented in rows. Unlike the TOWRE PDE, which gets increasingly more difficult, the Complex Nonsense Word Decoding measure maintains the same level of difficulty across the entire form. Only fully blended responses were coded as correct, yielding a score of correct nonsense words per minute. The average of the two forms was used in the analysis. The alternate form reliability obtained for this measure at pretest was .91 and .85 at posttest. The correlation between the two measures of decoding automaticity was .84 at pretest and .78 at posttest.

Text Reading Fluency

Three second-grade passages from DIBELS ORF matched for readability and topic were used to gather rate and accuracy data at pretest and posttest. Alternate-form reliabilities were .88 and .91 at pretest and .88 and .85 at posttest. Mean words read correctly across the three passages were used in the analysis. To monitor growth in text reading fluency across the intervention, students read a second-grade DIBELS passage every 10th session. Due to absences, some students were not assessed exactly at the 10th session; some were assessed after eight or nine sessions. This difference in assessment schedule was taken into account during the data analysis.

Reading Comprehension and Vocabulary

Reading comprehension was measured using the Reading Comprehension subtest of the WJ III (2006), an untimed cloze-type measure. Standard scores were determined using grade-level norms. The picture vocabulary subtest of the WJ III was used to measure receptive vocabulary knowledge for screening purposes only.

RESULTS

Descriptive Statistics

Table 2 shows the descriptive data for each of the variables for the entire sample and each intervention condition. Despite random assignment of participants to groups, there were differences between the two groups at pretest that favored the accuracy group ($d = .47$ for decoding accuracy, .59 for decoding automaticity, and .24 for oral reading fluency). None were significant except for the TOWRE PDE, on which the accuracy group began significantly higher than the automaticity group ($t = 2.08$, $df = 54$, $p = .042$, $d = .58$).

Data Analysis

In order to control for nesting effects of tutor and compare the effects of the two practice conditions on the decoding and reading comprehension outcomes, we used multilevel hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) with restricted maximum likelihood estimation to analyze the data. Student variables included the pretest covariate, posttest, intervention condition, word page reached and were entered at Level 1. Differences due to tutors were accounted for in Level 2.

A model was built for each student outcome (decoding accuracy, decoding automaticity, reading comprehension) and an exemplar model is included in Appendix B. The outcomes were posttest score controlling for pretest. Due to the nature of the intervention, which required mastery of material before moving forward, students reached different points in the curriculum. This led to the planned use of curriculum mastery as measured by the farthest word page reached as a covariate to control for different levels of

TABLE 2
Mean Raw Scores, Standard Scores, and (Standard Deviations) on Measures for Students at Pretest and Posttest by Group

	Entire Sample (N = 56)		*Accuracy (N = 27)*		*Automaticity (N = 29)*	
Measures	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
KTEA decoding accuracy						
Raw score	7.86 (6.78)	12.59 (9.09)	9.41 (6.28)	14.41 (9.58)	6.41 (7.02)	10.90 (8.42)
Standard score	90.75 (10.91)	97.79 (12.66)	93.59 (9.76)	100.07 (12.62)	88.10 (11.41)	95.66 (12.53)
TOWRE PDE						
Raw score	11.96 (5.51)	15.31 (6.59)	13.50 (5.16)	15.52 (5.39)	10.52 (5.53)	15.12 (7.64)
Standard score	92.55 (6.41)	96.45 (7.67)	94.37 (6.04)	96.63 (6.22)	90.86 (6.37)	96.28 (8.92)
Complex nonsense word decoding						
Raw score	7.05 (5.41)	10.14 (6.83)	8.48 (4.66)	10.07 (5.77)	5.72 (5.80)	10.21 (7.79)
DIBELS oral reading fluency						
Raw score	40.70 (17.24)	53.88 (19.51)	43.09 (15.98)	53.91 (18.01)	38.47 (18.33)	53.84 (21.12)
WJ III reading comprehension						
Raw score	16.41 (3.33)	18.84 (3.85)	16.59 (2.93)	19.89 (4.27)	16.24 (3.71)	17.82 (3.14)
Standard score	81.27 (9.42)	83.57 (9.85)	81.89 (8.42)	86.37 (10.63)	80.69 (10.38)	80.97 (8.42)
Sound page reached		13.98 (2.76)		15.30 (1.92)		12.80 (2.89)
Word page reached		12.26 (5.91)		16.63 (5.23)		8.33 (3.02)

exposure to the curriculum materials. The two covariates (pretest and farthest word page reached) were centered at the grand mean for the sample. In the exemplar model in Appendix B, the mean decoding accuracy is the fitted score for a child with a mean pretest score who reached the mean point in the curriculum in the accuracy group. All effect sizes are calculated as the difference between the adjusted posttest mean for the automaticity group minus the adjusted posttest mean for the accuracy group divided by the standard deviation of the accuracy group at pretest (Maxwell & Delaney, 2004).

To control for tutor effects and examine the effects of the two practice conditions on growth in the children's grade-level reading, we used HLM growth curve modeling to estimate the slope and intercept of text reading fluency while controlling for curriculum mastery. Reading fluency score and session number were entered into Level 1, student variables (e.g., intervention condition, word page reached) were entered at Level 2, and tutors were entered as Level 3. This model is included in Appendix B. The covariate was centered at the grand mean for the sample.

We next examined whether the two practice conditions differentially affected the progress of students with disabilities. To do this, we included disability status (0 = no disability, 1 = disability) and a Disability × Treatment interaction at Level 1 in the models previously described.

TABLE 3
Hierarchical Linear Modeling Results for Decoding Accuracy, Decoding Automaticity, and Reading Comprehension

Fixed Effect	*Coefficient*	SE	t(df)	p
Decoding Accuracy				
Intercept, β_0	12.95	0.95	13.64 (8)	<.001
Condition, β_1	2.45	1.89	1.30 (52)	.200
Word page, β_2	0.43	0.18	2.41 (52)	.020
Pretest, β_3	0.87	0.13	6.78 (52)	.001
Decoding Automaticity: TOWRE PDE				
Intercept, β_0	12.95	0.95	13.64 (8)	<.001
Condition, β_1	4.56	1.54	2.96 (52)	.005
Word page, β_2	0.28	0.16	1.75 (52)	.085
Pretest, β_3	0.87	0.13	6.78 (52)	<.001
Decoding Automaticity: Complex Nonsense Word Decoding				
Intercept, β_0	7.23	1.07	6.75 (8)	<.001
Condition, β_1	5.61	1.73	3.23 (52)	.003
Word page, β_2	0.39	0.18	2.18 (52)	.033
Pretest, β_3	0.79	0.14	.14 (52)	<.001
Reading Comprehension				
Intercept, β_0	18.94	0.78	24.24 (8)	<.001
Condition, β_1	−0.51	1.22	−0.42 (50)	.673
Word page, β_2	0.18	0.11	1.64 (50)	.107
Pretest, β_3	0.76	0.13	5.81 (50)	<.001

Note: Model deviance for decoding accuracy = 328.06, model deviance for TOWRE PDE = 307.96, model deviance for complex nonsense word decoding = 320.68, model deviance for reading comprehension = 248.81. TOWRE PDE = phonemic decoding efficiency of the *Test of Word Reading Efficiency*.

Comparison of Practice Conditions

We first examined the effects of practice condition on the decoding accuracy, decoding automaticity, and reading comprehension of the second-grade struggling students. As seen in Table 3, the pretest scores were significant for all outcomes, meaning that students who scored above the mean at the pretest also had higher scores at the posttest. There were no Pretest × Condition interactions, suggesting that the conditions were equally effective for children regardless of where they began.

Students in the two conditions did not differ on decoding accuracy (see Table 3). Students in the accuracy group who scored at the mean on the pretest and mastered a mean amount

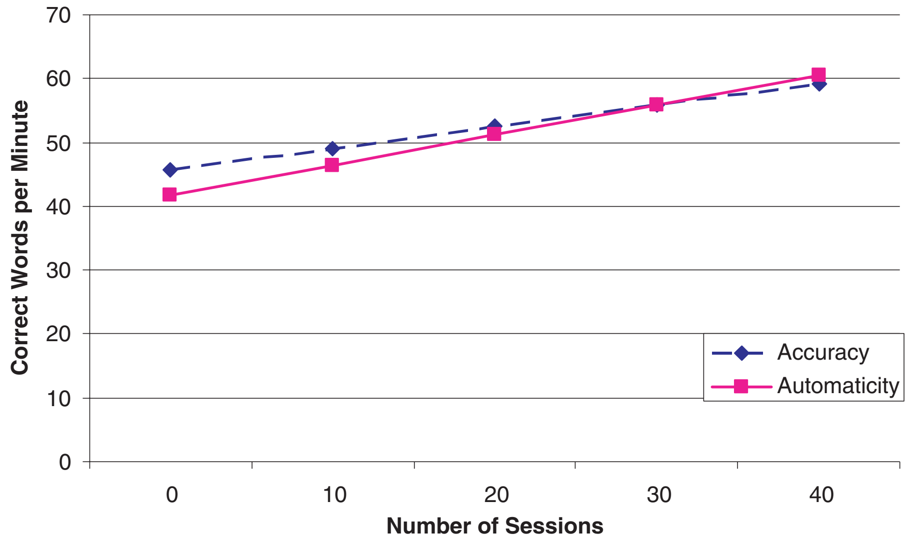

FIGURE 1 Growth in second-grade text reading fluency by practice condition for students at grand mean on pretest and at condition mean on curriculum mastery.
Note: The range of sessions completed by participants ranged from 17 to 40.

TABLE 4
Hierarchical Linear Growth Curve Modeling Results for Text Reading Fluency

Fixed Effect	*Coefficient*	SE	t(df)	p
Overall intercept at Pretest, π_0	34.99	4.02	8.71(8)	<.001
Condition, β_1	16.41	6.22	2.64 (53)	.011
Word page, β_2	2.45	.521	4.71 (53)	.020
Overall slope per session, π_1	0.34	.075	4.48 (264)	<.001
Condition, β_{11}	0.13	.130	1.10 (264)	.273
Word page, β_{12}	−0.001	.010	−0.08 (264)	.935

Note: Model deviance = 2068.09.

of curriculum had a predicted mean score of 11.08 while the fitted score for the students in the accuracy + automaticity group was 13.52 ($d = .39$).

There were significant differences favoring the accuracy + automaticity group on the two measures of decoding automaticity (see Table 2). For the TOWRE PDE, the accuracy + automaticity group at the mean on the pretest and amount of curriculum mastery had a predicted mean score 4.55 items higher than that of the accuracy group ($d = .88$). A similar pattern was found for the researcher-created Complex Nonsense Word Decoding. There was a 3.76-item advantage for the students in the accuracy + automaticity condition at the mean on the pretest and amount of curriculum mastery ($d = .80$). There were no differences between the groups on reading comprehension; the accuracy group at the mean on the pretest and curriculum mastery had a predicted score of 18.94 items with the accuracy + automaticity group was slightly lower (18.43, $d = -.17$).

As seen in Figure 1 and Table 4, while there was a significant amount of growth in reading grade-level connected text from the beginning to the end of the intervention, there were no differences between the practice conditions on the number of correct words per minute. On average, children in the accuracy group with an average amount of curriculum mastery grew .34 words per session while the children in the accuracy + automaticity group grew an average of .47 words per session.

To determine whether the students with disabilities responded differentially to the practice conditions, we tested whether a differential response (Disability × Treatment) emerged as a function of having a school-identified disability. No significant main effect due to disability or interaction with intervention was found for any outcome measure. In our small sample, students with disabilities responded as well as their nondisabled peers to the two practice conditions.

Comparison of Interventions

While we expected that students would reach different places in the curriculum due to the mastery nature of the intervention, we did not expect the large difference between the two groups that we found (see Table 2). On average, for the sound practice pages, children in the accuracy condition reached page 15 of a total of 16 possible pages while the children in the accuracy + automaticity condition reached page 13. While a small difference, it is nonetheless significant ($t = 3.87$, $df = 50.74$, $p < .001$). For the word pages, the difference is quite a bit larger. On average, students in the accuracy condition reached page 17 of 40 possible while the children in the accuracy + automaticity condition reached page 8 ($t = 7.22$, $df = 40.65$, $p < .001$).

The students in the accuracy group reached more than twice as far into the curriculum as the students in the accuracy + automaticity group and led us to ask another research question about the difference in overall effect of the two interventions. Given the limited resources of time and personnel available in schools, it stands to reason that the intervention that worked better overall would be preferable to one that did not. While our planned research questions about types of practice are theoretically interesting and important to answer, schools are not able to control for the amount of curriculum covered. So we examined our data a final time to answer the question of which intervention, one focused on accuracy or one focused on automaticity, is more effective in increasing the decoding accuracy, decoding automaticity, reading comprehension, and text reading fluency of the struggling students in our study. Using HLM, we fit a model for the outcomes with child-level data (e.g., pretest, posttest, condition) while controlling for differences due to tutor. Other than controlling for where children started, we used no covariate.

As seen in Table 5, we found the same pattern of results for decoding accuracy, decoding automaticity, and text reading fluency. There were no differences due to condition for decoding accuracy and text reading fluency. For the TOWRE PDE, students in the accuracy group at the mean for the pretest had a fitted posttest score of 13.96 while the students in the accuracy + automaticity group had a score of 16.57 ($d = .51$). Similarly, for Complex Nonsense Word Decoding, students who scored at the average on the pretest ended the intervention at 8.65 while those in the accuracy + automaticity condition had a fitted posttest score of 11.52 ($d = .61$). The advantage for the accuracy + automaticity group was much smaller when we considered the intervention holistically. When we examined reading comprehension, the advantage turns to the accuracy condition. On average, those who scored at the mean on the pretest in the accuracy condition scored 2.04 items more than children in the accuracy + automaticity condition ($d = -.70$).

TABLE 5
Hierarchical Linear Modeling Results for Decoding Accuracy, Decoding Automaticity, and Reading Comprehension Without Control for Curriculum Mastery

Fixed Effect	*Coefficient*	SE	*t*(*df*)	*p*
Decoding Accuracy				
Intercept, β_0	12.68	1.01	12.60	<.001
Condition, β_1	−0.38	1.36	−0.28	.778
Pretest, β_2	1.13	0.01	11.39	<.001
Decoding Automaticity: TOWRE PDE				
Intercept, β_0	13.96	0.77	18.13 (8)	<.001
Condition, β_1	2.62	1.09	2.40 (53)	.020
Pretest, β_2	1.10	0.10	10.13 (53)	<.001
Decoding Automaticity: Complex Nonsense Word Decoding				
Intercept, β_0	8.65	.88	9.82 (8)	.000
Condition, β_1	2.87	1.24	2.31 (53)	.025
Pretest, β_2	0.99	.12	8.59 (53)	<.001
Reading Comprehension				
Intercept, β_0	19.76	0.54	36.75 (8)	<.001
Condition, β_1	−2.04	0.69	−2.96 (51)	.005
Pretest, β_2	0.88	0.11	8.35 (51)	<.001

Note: Model deviance for decoding accuracy = 329.99, model deviance for TOWRE PDE = 307.35, model deviance for complex nonsense word decoding = 307.96, model deviance for reading comprehension = 246.84. TOWRE PDE = phonemic decoding efficiency of the *Test of Word Reading Efficiency*.

DISCUSSION

Students in both practice conditions improved in overall levels of performance during the repeated reading intervention even though no connected text was used in the intervention. There was no difference between the two groups on decoding accuracy, reading comprehension, or text reading fluency at the end of the study. The students in the accuracy + automaticity practice condition improved more than the students in the accuracy condition on both measures of decoding automaticity. There were no differential effects of treatment based on pretest scores or disability status; students appeared to benefit equally despite their initial status on the outcome variable or whether they had been identified with a disability.

Decoding Accuracy and Automaticity

We had not expected a large difference between the groups on the decoding accuracy measure because both groups were focused on that type of practice. However, there was a small effect favoring the accuracy + automaticity group (effect size = .39). Because automaticity in a skill tends to increase over time with practice (Binder et al., 2002), it would be reasonable to assume that, with more time in intervention, this trend in favor of the accuracy + automaticity practice condition would have solidified into a stronger effect.

A strong effect in favor of the accuracy + automaticity practice condition was found for both the TOWRE PDE, which began with easier nonsense words and became more difficult, and the Complex Nonsense Word Decoding, which had a consistent level of difficulty. Students who practiced letter sounds and word families to automaticity were more quick and accurate in pronouncing nonsense words than those who practiced to accuracy when curriculum mastery is controlled. Thus, we conclude that the automaticity developed with the isolated letter sounds and the word families led to the difference between the two groups, at least for students who reached the average place in the reading materials. This difference appeared even though there was little overlap between the practiced word families and the assessment measures. Both conditions produced improvement on this measure, which is consistent with previous research on transfer of learning (e.g., Lovett et al., 1994) that demonstrates the power of decoding instruction in developing skill for reading untaught words and patterns, but the advantage for the automaticity condition in this study suggests the need for further research to examine the role of automaticity development in skill transfer.

Text Reading Fluency

Despite the lack of differences between groups, the growth models show that both conditions of practice with isolated letter sounds and words led to increased text reading fluency in unpracticed grade-level connected text. The growth experienced by the two groups was comparable to national normative estimates of weekly average growth in second grade (Hasbrouck & Tindal, 2006). The accuracy group grew on average .34 words per session. Assuming most students had three sessions a week, that is a growth of 1.02 words per week, just below the growth for the 25th percentile of Hasbrouck and Tindal. The automaticity + accuracy group grew a little bit more, .47 words per session, or 1.41 words per week. While a small increase, it nonetheless put their growth just above that of all the second-graders in the Hasbrouck and Tindal sample.

Our results are consistent with the findings of some researchers (Fleisher et al., 1979; Levy et al., 1997; Tan & Nicholson, 1997) and contrasts with the findings of others (Dahl, 1979; Daly & Martens, 1994). For some students who are not fluent readers, practice in isolated letter sounds and within-word patterns appears to result in increased text reading fluency, even in grade-level text. However, we had hypothesized that the students in the condition focused on both rate and accuracy would grow faster than the students who merely focused on accuracy due to the theoretical increased value of automaticity (Laberge & Samuels, 1974), but we did not find such a difference. One possible reason for the lack of difference is the difficulty of the passages. Perhaps second-grade passages were still too difficult to capture differential amounts of growth. Haughton (1972) determined that a high rate of errors was linked with a low rate of correct responses. Therefore, the second-grade text, which was too challenging for most participants to read accurately, was unlikely to demonstrate differences in automaticity or reading rate. A set of passages at the first-grade level would have allowed for demonstration of a greater degree of accuracy and, consequently, would have made differences in automaticity more likely to be detectable. Another possible reason for the lack of differences in passage reading fluency is that, because most of the participants began the intervention quite inaccurate on both letter sounds and word families, it is possible that students in both groups were focused on accuracy at the beginning of the intervention because the students in the accuracy + automaticity condition received both accuracy and rate cues depending on their needs. In addition, the accuracy + automaticity condition began with a relatively low criterion of rate to move to the next page (25 in the word families), which was then increased as they moved through the materials. Previous findings indicate that a criterion substantially higher would be necessary to indicate automaticity (Haughton, 1972). Perhaps these two factors limited the differences between the two practice conditions.

Holistic Comparison of Interventions

Because we wanted to examine the effects of different types of practice, it was necessary to control the amount of curriculum mastery among the children. However, curriculum coverage became an important aspect to consider when evaluating the merits of these two interventions, especially from a school practitioner's perspective. Because schools have limited resources in terms of time and resources, it seemed prudent to evaluate the two interventions in terms of their overall effect on the students' reading, including the effect of differing amounts of curriculum coverage.

To examine this, we analyzed each outcome while controlling for only initial status on that outcome. This led to some interesting findings. One was that the large effect in favor of the automaticity group was lessened in both measures of decoding automaticity. The effect size went from .88 to .51 on the TOWRE PDE and from .80 to .61 on the Complex Nonsense Word Fluency. This was a substantial decrease in effectiveness that is probably due to the difference in exposure to the content.

Another finding was the change from no difference (effect size = −.17) between the two groups on reading comprehension to a large difference favoring the accuracy group (effect size = −.70). We think this change may be due to several factors. The first is that the accuracy group had been repeatedly cued to read carefully, a behavior that may have carried over to this measure and led to an advantage in the untimed context. It is also possible that the increased benefit may be due to the larger exposure to different word families found in the accuracy group, which may have helped them decode the unfamiliar words in the WJ III subtest. Linked to this is the notion that automaticity is not an all-or-nothing proposition; instead, it develops over a series of exposures in a predictable curvilinear pattern of increasing speed to an asymptote. Perhaps the exposure to more letter sounds and word patterns in the accuracy helped them become a bit more automatic than the children who had no opportunity at all to practice these items. According to Logan's (1988, 1997) instance theory of automatization, each encounter with an item increases the speed of access, meaning that even one encounter will make a reader grow in his or her automaticity. Because of the overlap of items built into the materials, students who made it further into the materials had several opportunities to encounter a particular letter sound or rime not available to those who did not.

Limitations

There were important limitations to this study that affect the interpretations of these results. The first is the relatively small sample size that limits generalization and did not allow us to take into account the larger nesting variables (e.g., classroom or school). Also, there was a problem with the accuracy condition that we did not foresee. In several of the accuracy groups, the children spontaneously began to compete to see who could read further in the text despite the tutor's cues to ignore rate and focus on reading accurately. This may have lessened the distinctions between the two practice conditions.

Another limitation was that students in the accuracy condition were higher performing at the start of the study than were students in the accuracy + automaticity condition. While we found only one of these differences to be statistically significant, the effect sizes for two of the four nonsignificant differences were substantial, making it difficult to ignore the possibility that the groups were not equivalent in some important ways. Although we covaried out pretest performance, this statistical adjustment does not rule out the possibility that the tasks were differentially appropriate for the two groups. We examined the demographics of the two groups (presented in

Table 1), and found the two groups were very similar in terms of ethnicity, number of English language learners, and number of students with disabilities. Where they differed was in the type of disability represented. The accuracy + automaticity group had students with emotional/behavioral disabilities, developmental disabilities, and two students with both specific learning disabilities (SLD) and speech/language disabilities, while the accuracy group had only students with SLD and health impairment, which may have made the groups not equivalent. The students were randomly assigned to condition, meaning that any nonequivalence happened by chance and was unavoidable.

Instructional group size and composition may also have played a role. The instructional groups ranged in size from two to four students and many of the students exhibited attentional and behavioral difficulties. The groups of four proved to be very challenging for some of the tutors to manage and reduced the amount of individual attention, modeling, and feedback provided to the students. Smaller group sizes or tutors more experienced in managing challenging behaviors may have made a difference in our results.

Due to challenges in the school schedule and a large number of child absences, it was difficult to control the number of sessions each child completed. While 40 sessions was the goal for all students, only 25 of the 56 completers received between 35 and 40 sessions and only 6 received the full complement. It is not clear in this study that the differences found are due to the practice conditions or differences in amount of instruction. We are also limited by the lack of a contact-control group. It is possible that some of the results were simply due to the increased attention children received in the small-group setting or maturation during the course of the study, although previous studies that have established positive effects of similar interventions when compared with a control condition indicate that this is unlikely.

CONCLUSIONS

Additional research to address these limitations and clarify the role practice to automaticity plays in intervention designed to increase text reading fluency is needed. Replication of this study with smaller groups of students, more sessions, and sufficient schools to allow for more rigorous statistical analysis is warranted. Also, it is not clear that the answer to our question about the most effective type of practice is one or the other. Perhaps the answer is rather one of timing of conditions. That is, perhaps an ideal scenario would be presenting initial cues focused on accuracy until a solid foundation has been reached before moving to cues focused on rate and accuracy. As Shuell (1990) suggests, "during the initial phase of learning the individual typically acquires isolated facts that are interpreted in terms of preexisting schemata and added to existing knowledge structures. Gradually, the learner begins to assemble these pieces into new schemata that provide him or her with more conceptual power until a level of automaticity is achieved" (p. 531). Many of the students in this study began with very inaccurate decoding skills, so perhaps a gradual approach would be more effective and help address the issue of differences in curriculum coverage, as well.

Fluency deficits caused by weak word-level skills pose a continuing challenge for teachers of developing readers. This study reinforces the notion that word-level practice improves word-level skills, both in lists and in connected text. Practice that emphasizes only accuracy allows students exposure to more text, which in turn seems to enhance comprehension while practice that emphasizes both accuracy and automaticity promotes the development of decoding efficiency. Both approaches have advantages, and the best instructional method may be a sequential combination of these approaches, beginning with a focus on accuracy and progressing to a focus on automaticity.

ACKNOWLEDGMENTS

The work presented in this article was supported by Grant H324N040039 from the U.S. Department of Education, Office of Special Education Programs. This article does not necessarily reflect the positions or policies of this funding agency and no official endorsement should be inferred.

We are grateful to Maren Anderson, Shelly Levy, Betty Lopez, Nikole Duvalle, Melissa Blanco, Cady Garey, Wanda Yachnis, Malina Mendez, and Anna Lee for their work as intervention teachers and to Marriah Vinson and Noa Hannah for their help with data collection. We appreciate Robert Abbot for his help with data analysis and Joseph Jenkins for his helpful comments on an earlier draft of this manuscript. We especially thank the children, parents, teachers, and principals in Washington and Florida who made this research possible.

REFERENCES

Berninger, V. W., Abbott, R. D., Vermeulen, K., & Fulton, C. M. (2006). Paths to reading comprehension in at-risk second-grade readers. *Journal of Learning Disabilities, 39*, 334–351.

Binder, C., Haughton, E., & Bateman, B. (2002). *Fluency: Achieving true mastery in the learning process*. Professional Papers in Special Education. University of Virginia Curry School of Special Education.

Blachman, B. A., Ball, E. W., Black, R., & Tangel, D. M. (2000). *Road to the code: A phonological awareness program for young children*. Baltimore: Brookes.

Brown, G. D. A., & Deavers, R. P. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Psychology, 73*, 208–242.

Bryan, W. L., & Harter, N. (1899). Studies of the telegraphic language. The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345–375.

Bucklin, B. R., Dickinson, A. M., & Brethower, D. M. (2000). A comparison of the effects of fluency training and accuracy training on application and retention. *Performance Improvement Quarterly, 13(3)*, 140–163.

Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35*, 386–406.

Dahl, P. R. (1979). An experimental program for teaching high speed word recognition and comprehension skills. In J. E. Button, T. Lovitt, & T. Rowland (Eds.), *Communications research in learning disabilities and mental retardation* (pp. 33–65). Baltimore: University Park Press.

Daly, E. J., & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis, 27*, 459–469.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.

Ehri, L. C. (2002). Phases of acquisition in learning to read words and implications for teaching. In R. Stainthorp & P. Tomlinson (Eds.),

Learning and teaching reading. London: British Journal of Educational Psychology Monograph Series II.
Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. *Reading Research Quarterly, 15*, 30–48.
Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading disabilities in at-risk children. *Journal of Educational Psychology, 90*, 37–55.
Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
Goswami, U., & Mead, F. (1992). Onset and rime awareness and analogies in reading. *Reading Research Quarterly, 27*, 152–162.
Hasbrouck, J., & Tindal, G.A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636–644.
Haughton, E. C. (1972). Aims: Growing and sharing. In J. B. Jordan & L. S. Robbins (Eds.), *Let's try doing something else kind of thing.* Arlington, VA: Council on Exceptional Children.
Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading & Writing Quarterly, 25*(1), 4–32.
Hudson, R. F., Torgesen, J. K., & Schatschneider, C. (2006). *Relative effects of speed and accuracy cues on the oral reading fluency of young children*. Unpublished manuscript. Florida Center for Reading Research, Tallahassee.
Kaufman, A. S., & Kaufman, N. L. (2004). *KTEA II: Kaufman test of educational achievement comprehensive.* Circle Pines, MN: American Guidance Services.
Kessler, B., & Treiman, R. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology, 24*, 267–289.
LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.
LeVasseur, V. M., Macaruso, P., & Shankweiler, D. (2008). Promoting gains in reading fluency: A comparison of three approaches. *Reading and Writing: An Interdisciplinary Journal, 21*, 205–230.
Levy, B. A. (2001). Moving the bottom. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 357–379). Timonium, MD: York.
Levy, B. A., Abello, B., & Lysynchuk, L. (1997). Transfer from word training to reading in context: Gains in reading fluency and comprehension. *Learning Disabilities Quarterly, 20*, 173–188.
Levy, B. A., & Lysynchuk, L. (1997). Beginning word recognition: Benefits of training by segmentation and whole word methods. *Scientific Studies of Reading, 1*(4), 359–387.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.
Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 13,* 123–146.
Lovett, M., Borden, S., DeLuca, T., Lacerenza, L., Bensen, N., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. *Developmental Psychology, 30*, 805–822.
Martin-Chang, S. L., & Levy, B. A. (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing, 18*, 343–376.
Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Erlbaum.
McKay, M. F., & Thompson, G. B. (2009). Reading vocabulary influences in phonological recoding during the development of reading skill: A re-examination of theory and practice. *Reading and Writing, 22*, 167–184.
Moats, L. C. (2000). *Speech to print: Language essentials for teachers*. Baltimore: Brookes.
Morgan, P. L., & Sideridis, G. D. (2006). Contrasting the effectiveness of fluency interventions for students with or at risk for learning disabilities: A multilevel random coefficient modeling meta-analysis. *Learning Disabilities: Research & Practice, 21*, 191–210.
National Reading Panel. (2000). *Report of the National Reading Panel: Reports of the subgroups*. Washington, DC: U.S. Department of Health and Human Services, National Institute of Health.
O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1987). The effects of repeated readings and attentional cues on the reading fluency and comprehension of learning disabled readers. *Learning Disabilities Research, 2*, 103–109.
Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.
Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*, 461–469.
Pressley, M., Hilden, K., & Shankland, R. (2006). *An evaluation of end-grade-3 Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little*. Retrieved August 12, 2006, from http://www.msularc.org/docu/dibels_submitted.pdf
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
Samuels, S. J., & Flor, R. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly, 13*(2), 107–122.
Shuell, T. J. (1990). Phases of meaningful learning. *Review of Educational Research, 60*, 531–547.
Tan, A., & Nicholson, T. (1997). Flashcards revisited: Training poor readers to read words faster improves their comprehension of text. *Journal of Educational Psychology*, *89*, 276–288.
Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education, 25*, 252–261.
Torgesen, J. K. (1986). Computers and cognition in reading: A focus on decoding fluency. *Exceptional Children, 53*(2), 157–162.
Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: Critical factors for struggling readers. In S.J. Samuels & J. Farstrup (Eds.), *What research has to say about fluency instruction*. Newark, DE: International Reading Association.
Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). Parkton, MD: York.
Torgesen, J. K., Wagner, R. K., & Rashotte, C. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.
Walton, P. D., & Walton, L.M. (2002). Beginning reading by teaching in rime analogy: Effects on phonological skills, letter-sound knowledge, working memory, and word-reading strategies. *Scientific Studies of Reading, 6*(1), 79–115.
Woodcock, R. W., McGrew, K. S., & Mather, M. (2001). *Woodcock-Johnson III tests of academic achievement*. Itasca, IL: Riverside.

APPENDIX A

Letter Sound Progression

Page 1	m, s, r, f, short a
Page 2	p, c, t, l, short i
Page 3	g, d, k, n, short o
Page 4	h, b, j, z, short u
Page 5	v, y, w, x, short e
Page 6	Mastery test single letters and short vowels
Page 7	qu, th, ch, sh
Page 8	or, ing, ar, ir
Page 9	wh, ur, er, ow (long o)
Page 10	Review
Page 11	oa, ee, ea, ai
Page 12	ay, igh, oo, oi
Page 13	gh, ow, wr, ew
Page 14	ph, kn, ui, ou
Page 15	Review
Page 16	Mastery test of all taught sounds

Word Family Progression

Page 1	at, am, as	Page 21	orn, or/ord, ore, ir/irl
Page 2	ap, im, it,	Page 22	irt, urn, url
Page 3	ag, in, ip	Page 23	Review
Page 4	ad, ill, id, og	Page 24	Mastery test: long vwl & r-controlled
Page 5	ab, od, ot, um	Page 25	ash, ath, amp, ang
Page 6	ob, op, oss, un	Page 26	ank, int, ink, inch
Page 7	ack, ug, ush, ud	Page 27	unk, unch, ump, elt
Page 8	ick, ock, ut, ell	Page 28	est, ent, end, ow
Page 9	uck, ed, et, eck	Page 29	own, owl, out, ound
Page 10	Review	Page 30	ouch, udge, ouse, ung
Page 11	Mastery Test: Short vowels	Page 31	ew, oon, oop, oot,
Page 12	open syllables	Page 32	oy, oil, ow (long o)
Page 13	ake, ame, ade, ate	Page 33	Review
Page 14	ine, ime, ite, ice	Page 34	Mastery test: single syllables
Page 15	ose, oke, ope, ue	Page 35	Syllable types: closed
Page 16	ute, eed, eet, een	Page 36	Syllable types: open
Page 17	eek, each, eal, ean	Page 37	Syllable types: V_e
Page 18	eak, eam, ing, ight	Page 38	Syllable types: -le
Page 19	ail, ain, ay, oat	Page 39	Syllable types: vowel team
Page 20	oal, ar, art, ard	Page 40	Overall mastery test

APPENDIX B

Exemplar Hierarchical Linear Model: Decoding Accuracy

Level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{condition})_{ij} + \beta_{2j}(\text{word page})_{ij} + \beta_{3j}(\text{pretest})_{ij} + R_{ij}$$

Level 2

$\beta_{0j} = \gamma_{00} + U_0$
$\beta_{1j=}\ \gamma_{10}$
$\beta_{2j=}\ \gamma_{20}$
$\beta_{3j=}\ \gamma_{30}$

Exemplar Hierarchical Linear Model: Text Reading Fluency

Level 1

$$Y_{ij} = \pi_0 + \pi_1(\text{session}) + e$$

Level 2

$\pi_0 = \beta_{00j} + \beta_{01j}(\text{condition})_{ij} + \beta_{02j}(\text{word page})_{ij} + R_{1ij}$
$\pi_1 = \beta_{10j} + \beta_{11j}(\text{condition})_{ij} + \beta_{12j}(\text{word page})_{ij}$

Level 3

$\beta_{00j} = \gamma_{000} + \mu_0$
$\beta_{01j=}\ \gamma_{010}$
$\beta_{02j=}\ \gamma_{020}$
$\beta_{10j=}\ \gamma_{100} + \mu_1$
$\beta_{11j=}\ \gamma_{110}$
$\beta_{12j} = \gamma_{120}$
$\beta_{3j=}\ \gamma_{100}$

About the Authors

Roxanne F. Hudson is an associate professor in Special Education at the University of Washington. Her research interests focus on the development of reading fluency in young children, interventions for struggling readers, literacy instruction for children on the Autism spectrum, and teacher education.

Holly B. Lane is an associate professor in Special Education at the University of Florida. Her research interests focus on reading acquisition and the prevention of reading disabilities through effective early instruction, intervention for struggling readers, teacher education; and learning disabilities across the life span.

Stephanie Arriazza-Allen is a recent graduate of the University of Florida Special Education PhD program. Her research interests are in early reading intervention for Spanish-speaking young children.

Carole Isakson is a project coordinator in Special Education at the University of Washington. Her interests are in teacher education, induction support for beginning teachers, and the use of assistive technology to facilitate inclusion of people with disabilities.

Taylor Richman is a PhD student in Multicultural Education at the University of Washington. His research interests are focused on the relations among race, ethnicity, and disability.