Orton-Gillingham and Orton-Gillingham–Based Reading Instruction:

A Review of the Literature

Kristen D. Ritchey
University of Delaware
Jennifer L. Goeke
Montclair State University

Orton-Gillingham (OG) and Orton-Gillingham–based reading instructional programs are commonly implemented reading programs in the United States. Twelve studies that employed quasi-experimental or experimental designs are reviewed. These studies included elementary students, adolescents, and college students. Of the 12 studies, 5 reported that the OG instruction was more effective than were comparison or control interventions for all measured outcomes, 4 reported that the OG instruction was more effective for at least 1 (but not all) outcomes in comparison to other intervention(s), 2 reported that the alternate instruction was more effective than the OG instruction, and 1 reported no significant differences once covariates were included. The largest effects were reported for word attack and nonword reading outcomes, with mean effect size of .82, and comprehension outcomes, with a mean effect size of .76. Following a summary of each study, limitations of the current review and implications for future research and practice are described. Given the small number of studies, the lack of methodological rigor of the existing studies, and the inconclusive findings of the effectiveness of OG programs, additional research is needed before the scientific basis can be established.

In the early part of the 20th century, the field of learning disabilities was still in its infancy. At that time, Dr. Samuel Orton, a child neurologist, estimated that approximately 10% of the school population had reading disabilities and speculated that reading problems were caused by neurological deficits. Orton's early conclusions formed two major tenets of the characteristics of dyslexia: (a) reading problems were caused by *strephosymbolia* (or twisted symbols), and (b) a lack of brain hemisphere dominance caused information to be processed in both hemispheres, resulting in mirror images (e.g., reversals of the letters *b* and *d*).

Although Orton's speculated causes of reading disabilities have not stood the test of time (phonological processing problems are currently accepted as the primary cause of reading disabilities), his ideas about appropriate reading instruction have. Specifically, Orton (1937) stated that the instructional approach should

> attempt to capitalize [*sic*] their [students'] auditory competence by teaching them the phonetic equivalents of the printed letters and the process of blending sequences of such equivalents so that they might be able to produce for themselves the spoken form of the word from its graphic counterpart. (p. 159)

The instructional approach conceived by Orton was developed into a curriculum by Anna Gillingham and Bessie Stillman and first described in the manual *Remedial Training for Children with Specific Disability in Reading, Spelling, and Penmanship* (1960). Currently in its eighth edition (Gillingham & Stillman, 1997), the manual remains the backbone of Orton-Gillingham (OG) instructional programs and other instructional programs derived from the original curriculum.

Characteristics of OG-Based Instruction

The OG approach is a systematic, sequential, multisensory, synthetic and phonics-based approach to teaching reading. Explicit instruction is provided in phonology and phonological awareness, sound-symbol correspondence, syllables, morphology, syntax, and semantics. A key characteristic of OG reading instruction is that it is multisensory, involving visual, auditory, and kinesthetic/tactile learning pathways, often referred to as the *Language Triangle*. Instruction directly teaches language components, is provided systematically and cumulatively, requires mastery and overlearning before students advance to new components, is based on ongoing diagnostic information and assessment, and is individualized to

Address: Kristen D. Ritchey, PhD, School of Education, Willard Hall University of Delaware, Newark, DE 19716; e-mail: kritchey@udel.edu; ph: (302) 831-2325

the needs of each student (Clark & Uhry, 1995). As a final point, OG instruction is to be provided by trained and qualified teachers, tutors, or other specialists.

Related reading programs have evolved from the original teaching methodology. Several authors worked directly with Orton, Gillingham, or Stillman to make their adaptations. Other adaptations were developed based on specific needs, such as modifying the approach to address whole class instruction or to meet the needs of adult learners. Well-known adaptations and extensions include Alphabetic Phonics (Cox, 1992), Wilson Reading System (Wilson, 1996), The Herman Method (Herman, 1993), Project ASSIST (Biasotto, 1993), The Slingerland Approach (Slingerland & Aho, 1994–1996), The Spalding Method (Spalding & Spalding, 1990), Starting Over (Knight, 1995), and Project Read (Enfield & Greene, 1997). Although this is not an exhaustive list, the programs listed above include materials, training, and research or evaluation. Across curricula, the underlying instructional principles are consistent with the original OG methodology, although the targeted age group, instructional setting, materials, or other differences may exist.

Need for Evidence From Scientifically Based Research

Since their development, OG and OG-based reading instruction have been commonly accepted and frequently delivered interventions for students with reading disabilities. Although the standard of identifying and using the best instructional practices is customary to special education, it appears that the widespread use of OG instruction has been fueled by anecdotal evidence and personal experience. It is possible, for example, that in comparison to general education reading instruction practices, other commonly used basal reading programs, and/or teachers' "eclectic" attempts at addressing reading difficulties, the relative effectiveness of OG experienced by practitioners is overwhelmingly positive. As with any intervention, however, adequate scientifically based evidence of the effectiveness of OG and OG-based programs should be sought before such instruction is provided for students.

Two current legal mandates require a scientific evidence base for reading instruction. First, No Child Left Behind (NCLB; 2002) legislation, which includes the Reading First initiatives, requires that schools employ "a learning system or program of reading instruction based on scientifically-based reading research" (§ 1202 [c][7][A]). New federal requirements define scientifically-based reading research as research that "applies rigorous, systematic and objective procedures to obtain valid knowledge relevant to reading development, reading instruction, and reading difficulties" (No Child Left Behind, 2002, Part B, Subpart 1, Section 1208(6)(A)). Second, OG or OG-based reading instruction often is specifically requested by parents and professionals for students receiving special education services. Bhat, Rapport, and Griffin (2000) reviewed 27 legal decisions based on due process cases in which parents requested specific reading methods. In 15 of the 27 cases (decided between 1989 and 1998), parents requested OG or a specific OG-based program (e.g., The Slingerland Approach, 1994–1996, or the Wilson Reading System, 1996). In an additional 4 cases, a language-based and/or multisensory program was requested.

Given the number of children who receive OG and OG-based reading instruction, as well as the current legal climate, it is important to establish whether there is, at this time, sufficient evidence to suggest that OG and OG-based reading instruction programs meet the requirements of scientifically-based reading instruction. The purpose of this paper is to review the literature on OG and OG-based reading instruction programs in comparison to other instructional approaches. Following a summary of research studies, methodological issues that confound understanding of the effectiveness of OG and OG-based programs will be described. Finally, implications for further research and instruction are discussed.

Method

To locate articles to include in this review, we searched the ERIC, ECER, and PsychInfo databases using the following key words: *Orton Gillingham, multisensory, reading, decoding,* and *phonics,* as well as the titles and authors of all OG-based curricula. We also searched the ProQuest dissertation database using the same terms and beginning with 1980. The reference lists of identified articles were examined for additional references. Finally, a hand search was conducted of the following journals: *Annals of Dyslexia* (titled *The Bulletin of the Orton Society* until 1981), *Journal of Learning Disabilities*, *Learning Disabilities Quarterly, Exceptional Children,* and *Journal of Special Education,* beginning with the earliest available issue through 2005.

Studies were selected for inclusion if they: (a) were published in a peer-refereed journal or were a doctoral dissertation, (b) studied a multisensory reading approach based on OG methods or philosophy, (c) were conducted using experimental, quasi-experimental, or single-subject research designs, and (d) included a sample size of at least 10 participants per experimental condition. Descriptive studies, pilot studies, program evaluations, and studies employing only a pretest–posttest design (without a control or comparison treatment) were excluded (e.g., Hutcheson, Selig, & Young, 1990; Ogden, Hindman, & Turner, 1989; Vickery, Reynolds, & Cochran, 1987). Three additional studies that applied OG principles to foreign language teaching were excluded (Sparks & Ganschow, 1993; Sparks, Ganschow, Kenneweg, & Miller, 1991; Sparks, Ganschow, Pohlman, Skinner, & Artzer, 1992). A corpus of 12 studies met the criteria for inclusion in this review.

All studies were coded for pertinent information, including participant characteristics, setting, experimental design, treatment conditions, comparison and/or control conditions,

dependent variables, and results. Each author coded each study, and interrater agreement was calculated using the following formula: Number of Agreements ÷ Number of Agreements + Disagreements. Interrater agreement across categories exceeded 99%.

Information for calculating effect sizes also was coded from each study. To determine effect sizes (*d*, Cohen, 1988), we identified and then used a spreadsheet to record the mean, the standard deviations, and the number of participants for treatment and control/comparison groups. We calculated effect sizes using the following formula: Mean of OG Treatment Group – Mean of Comparison/Control Group ÷ Pooled *SD*. When sufficient descriptive statistics were unavailable, effect sizes were estimated from *F* or *t* statistics, using the formula provided by Rosenthal (1991). When there were more than two treatment groups, effect sizes were calculated for both treatment groups. Interrater agreement for effect sizes exceeded 99%. Effect sizes can be interpreted as $d = .2$ as a small effect, $d = .5$ as a medium effect, and $d = .8$ as a large effect (Cohen, 1988).

Results

Twelve studies (ten articles and two dissertations) reporting the results of experimental or quasi-experimental research studies met the inclusion criteria. Table 1 provides descriptive information for each study, including (a) OG or OG-based instructional program, (b) comparison or control treatment(s), (c) experimental design, (d) participants, (e) setting, (f) outcome measures, (g) results, and (h) effect sizes. The results of the reviewed studies are presented by first describing each study and then summarizing the results across outcome measures and overall findings.

Elementary Students

Studies that investigated the effectiveness of OG approaches with elementary students varied in terms of how the approach was applied. In some cases, the OG approach was the primary method of reading instruction in general education classrooms. In others, the approach was used as an intervention method for students who were considered to be at risk for reading failure or as a secondary intervention for students with reading disabilities. The results of studies conducted with elementary students are organized according to instructional setting; first, the results of school-based interventions are reported, followed by clinic-based intervention studies.

School-Based Interventions. Stoner (1991) investigated the effectiveness of Project Read (Enfield & Greene, 1997) in a quasi-experimental design that compared Project Read to traditional basal reading instruction for first-, second-, and third-graders who were considered at risk for reading problems. Teachers implemented instruction in general education classrooms, and the progress of students was compared to students who had received traditional basal instruction the previous year (previous cohort). For first-grade children, there were significant differences on the *Stanford Achievement Test* in word study ($d = 1.15$), in word reading ($d = 1.06$), in comprehension ($d = .93$), and in total reading score ($d = 1.15$). All differences favored the Project Read intervention. For second- and third-grade children, there were no significant differences on any of the outcome measures. An additional analysis examined the results for children in classrooms where the teacher taught in the control condition the previous year and Project Read in the second year. We do not describe those results here, as the analysis duplicates the finding of the entire sample.

Litcher and Roberge (1979) investigated OG instruction as an early intervention program for first-grade students identified as at risk for reading problems by screening measures given to approximately 600 students in the school district. Twenty students were identified for each of 3 years and assigned to experimental interventions by matched pairs. Each year, students were taught OG reading and language instruction 3 hours a day and were compared to a group of students receiving traditional basal instruction in matched schools. At the end of first grade, students who received the OG instruction performed significantly higher than did the comparison group on the subtest and total scores of the *Gates MacGinitie Reading Test* (MacGinitie & MacGinitie, 1989) and the *Metropolitan Achievement Test*. Effect sizes for each year ranged from .75 to 2.88, indicating that the OG program produced greater performance (calculated from raw scores) than did traditional basal instruction for readers at risk for reading problems.

Foorman et al. (1997) also investigated the effects of Cox's (1992) Alphabetic Phonics (a synthetic phonics instructional program) for second- and third-grade students with reading disabilities in a quasi-experimental design. Two other instructional methods were used as comparison treatments: a sight word program (Edmark Reading Program, 1984) and an analytic phonics program (a modification of *Recipe for Reading;* Traub, & Bloom, 1992) that taught word identification at the onset-rime level). Students received one of the three types of reading instruction for 60 min per day across the school year and were assessed four times during the intervention. Growth curve analysis was used to model growth in phonological processing, orthographic processing, and word reading and to test for treatment effects. For phonological processing, the synthetic phonics instruction (OG) group significantly outperformed the analytic phonics instruction group ($d = .37$) and sight word instruction groups ($d = .59$) when controlling for age. When other covariates (gender, socioeconomic status, ethnicity, and verbal IQ) were controlled for, the synthetic phonics (OG) groups and analytic phonics groups no longer differed, although the synthetic group outperformed the sight word instruction group. Verbal IQ was a significant correlate. For orthographic processing, there were effects for age (older students

TABLE 1. Overview of 12 Studies of Orton-Gillingham (OG) or OG-Based Reading Instruction vs. Comparison/Control Interventions

Study	Instructional Program	Comparison/ control intervention	Experimental design	Participants	N[a]	Setting	Outcome measures	Results	Effect sizes
Chandler, Munday, Tunnell, & Windham (1993)	Alphabetic Phonics (Cox, 1992)	Traditional developmental reading course	Quasi-experimental	Community college students	43	Community college	NDRT	Alphabet Phonics > traditional instruction	d^* = −1.18
Dooley (1994)[b]	Alphabetic Phonics (Cox, 1992) combined with cooperative learning (MIRC) groups	Traditional middle school reading instruction	Quasi-experimental	Seventh-grade students in remedial reading course	151	Reading improvement classes	• SDRT–Word Attack • SDRT–Reading Rate • TORC– General Vocabulary • TORC–Syntactic Similarities • TORC–Paragraph Reading • TORC–Sentence Sequencing • TOWL–Thematic Maturity • TOWL–Contextual Style	MIRC > traditional instruction for all outcomes except TORC–General Vocabulary	• Word Attack: d = .46 • Reading Rate: d = .83 • General Vocabulary: d = .01 • Syntactic Similarities: d = .44 • Paragraph Reading: d = .42 • Sentence Sequencing: d = .51 *M* (comp): d = .46 • Thematic Maturity: d = 1.15 • Contextual Style: d = .79
Foorman et al. (1997)	Alphabetic Phonics (Cox, 1992; synthetic phonics instruction)	(a) Analytic phonics instruction (b) Sight word reading instruction	Quasi-experimental	Second- and third-grade students with reading disabilities	114	Special education resource rooms	• Phonological processing (synthesis & analysis) • Orthographic processing • WJ-R Basic Reading	When controlling for age, Alphabet Phonics instruction significantly outperformed analytic phonics instruction for phonological processing, orthographic processing, and word reading; Alphabet Phonics instruction was superior to sight word instruction for phonological processing and word reading, but not for orthographic processing; when additional covariates were controlled for, there were no longer significant differences for instruction groups	(without covariates) • Phonological Processing: OG > analytic phonics d = .37 OG > sight word d = .59 • Orthographic Processing: OG > analytic phonics d = .29 OG *ns* sight word d = .05 • Word Reading: OG > analytic phonics d = .39 OG > sight word d = .17

(Table continues)

(Table 1 continued)

Study	Instructional Program	Comparison/ control intervention	Experimental design	Participants	N[a]	Setting	Outcome measures	Results	Effect sizes
Guyer & Sabatino (1989)	OG	(a) Nonphonetic reading instruction (b) No intervention control	Quasi-experimental	College students	30	College	• WRAT–Reading • WRMT–Total Reading	Significant improvements from pretest to posttest for two of three groups; students in OG group made more growth than did other two groups	—
Guyer, Banks, & Guyer (1993)	Wilson Reading System (Wilson, 1996)	(a) Nonphonetic spelling instruction (b) No intervention control	Experimental design with random assignment to intervention conditions with additional comparison group receiving no intervention	College students	30	College	WRAT–Spelling	Wilson Reading System > nonphonetic, control; Wilson Reading System growth > nonphonetic, control growth	—
Hook, Macaruso, & Jones (2001)	OG	Fast ForWord (1998)	Quasi-experimental	Ages 7–12 yrs	31	Summer clinic	• WRMT–Word Identification • WRMT–Word Attack • LAC	• Word Identification:-*ns* growth • Word Attack: OG growth > Fast For Word growth • LAC: significant growth for both OG & Fast ForWord	• Word Identification: OG > Fast ForWord $d = .07$ • Word Attack: OG > Fast ForWord $d = .77$ • LAC: OG > FFW $d = .11$
Joshi, Dahlgren, & Boulware-Gooden (2002)	Language Basics: Elementary (Curriculum based on Alphabetic Phonics, Cox, 1992)	Houghton Mifflin basal series	Quasi-experimental	First-grade students	56	General education classrooms	• TOPA • WRMT–Word Attack • GM–Comprehension	• TOPA: Alphabet Phonics > basal • Word Attack: Alphabet Phonics > basal • Comprehension: Alphabet Phonics > basal	• TOPA: $d = .44$ • Word Attack: $d = 1.08$ • Comprehension: $d = .81$
Litcher & Roberge(1979)	OG	Basal reading instruction	Quasi-experimental	First-grade students at risk for reading problems	40	Special classrooms (experimental condition)	• MAT–Word knowledge • MAT–Word analysis • MAT– Comprehension • MAT–Total reading score • GM–Vocabulary • GM–Comprehension	• MAT all scores: OG > basal • GM all scores: OG > OG > basal	• Word Knowledge: $d = 1.13$ (Yr 1) $d = 1.79$ (Yr 2) $d = 1.75$ (Yr 3) $M\ d = 1.56$ • Word Analysis: $d = 1.37$ (Yr 1) $d = 1.56$ (Yr 2) $d = 2.88$ (Yr 3) $M\ d = 1.94$

(Table continues)

(Table 1 continued)

Study	Instructional Program	Comparison/ control intervention	Experimental design	Participants	N[a]	Setting	Outcome measures	Results	Effect sizes
									• Comprehension (MAT): *d* = .75 (Yr 1) *d* = 1.38 (Yr 2) *d* = 2.29 (Yr 3) *M d* = 1.47 • Total Reading score: *d* = .86 (Yr 1) *d* = 1.70 (Yr 2) *d* = 2.22 (Yr 3) *M d* = 1.59 • Vocabulary: *d* = 1.61 (Yr 1) *d* = 1.36 (Yr 2) *d* = 1.83 (Yr 3) *M d* = 1.60 • Comprehension (GM): *d* = 1.56 (Yr 1) *d* = 1.48 (Yr 2) *d* = 2.01 (Yr 3) *M d* = 1.68
Oakland, Black, Stanford, Nussbaum, & Balise (1998)	Dyslexia Training Program (no reference—Oakland et al would be the appropriate reference teacher- and video-directed versions)	Traditional remedial instruction	Quasi-experimental	Mean age = 11 yrs (*SD* = 2)	48	Clinical setting	• GM–Comprehension • WRAT–Reading • WRAT–Spelling • Decoding Skills Test	Both teacher-directed and video-directed programs yielded similar growth • Comprehension: Dyslexia Training Program > control • Word Reading: Dyslexia Training Program > control • Spelling: *ns* • Decoding–Monosyllabic Words: *ns* • Decoding–Polysyllabic Words: Dyslexia Training Program > control	• Comprehension: *d** = .65 • Word Reading: *d** = .73 • Spelling: *d** = .24 • Monosyllabic Words: *d** = .45 • Polysyllabic Words: *d** = .80
Simpson, Swanson, & Kunkel (1992)	OG	Remedial English instruction	Quasi-experimental	13–18 yrs (boys)	63	Juvenile detention facility	WRMT	OG growth > traditional class growth Students receiving more than 50 hours of instruction made more growth	*d* = .71

(Table continues)

(Table 1 continued)

Study	Instructional Program	Comparison/ control intervention	Experimental design	Participants	*N*[a]	Setting	Outcome measures	Results	Effect sizes
Stoner (1991)	Project Read (Enfield & Greene, 1997)	Basal reading instruction	Quasi-experimental with control students enrolled in grade during the previous academic year	First-, second-, and third-grade children (at risk for reading problems)	Gr. 1 *n* = 130 Gr. 2 *n* = 70 Gr. 3 *n* = 83	General education classrooms	• SAT–Word Study • SAT–Word Reading • SAT–Comprehension • SAT–Total Reading score	Full sample Gr. 1: OG > basal for all outcomes Gr. 2: *ns* for any outcome Gr. 3: *ns* for any outcome	• Word Study *d* = 1.15 (Gr 1) *d* = –.04 (Gr 2) *M d* = .56 • Word Reading *d* = 1.06 (Gr 1) *d* = –.21 (Gr 2 *d* = .31 (Gr 3) *M d* = .39 • Comprehension *d* = .93 (Gr 1) *d* = –.40 (Gr 2) *d* = .02 (Gr 3) *M d* = .18 • Total Reading score *d* = 1.15 (Gr 1) *d* = –.30 (Gr 2) *d* = .17 (Gr 3) mean *d* = .34
Westrich-Bond (1993)[b]	OG	Basal reading instruction (Ginn, no reference available)	Quasi-experimental	Students 6–12 years old with learning disabilities	72	Resource room or self-contained special education classroom	• WRMT–Word Identification • WRMT–Word Attack	No significant differences for type of reading instruction; there were differences in reading gains across classroom setting	• Word Identification Resource room *d* = –.40 Self-contained special education classroom *d* = –1.40 • Word Attack Resource room *d* = .07 Self-contained special education classroom *d* = –.03

Note. Dash indicates insufficient statistics to calculate effect sizes; *d**= effect size calculated from *t* or *F* statistic. GM = *Gates MacGinitie Reading Test* (MacGinitie & MacGinitie, 1989); LAC = *Lindamood Auditory Conceptualization Test* (Lindamood & Lindamood, 1979); MIRC = Multisensory integrated reading and composition; OG = Orton Gillingham reading instruction; NTRT = *Nelson Denney Reading Test;* SAT = *Stanford Achievement Test;* SDRT = *Stanford Diagnostic Reading Test* (Karlsen & Gardner, 1985); TOPA = *Test of Phonological Awareness* (Torgesen & Bryant, 1994); TORC = *Test of Reading Comprehension* (Brown, Hammill, & Wiederholt, 1978); TOWL = *Test of Written Language* (Hammill & Larsen, 1988); WRMT = *Woodcock Reading Mastery Test* (Woodcock, 1973, 1989, or 1998); WRAT = *Wide Range Achievement Test* (Jastak & Wilkinson, 1984); WJ–R = *Woodcock Johnson Psycho-Educational Battery–Revised* (Woodcock & Johnson, 1989).
[a]*N* indicates number of participants after accounting for attrition. [b]Indicates an unpublished dissertation.

had higher end-of-year performance and rates of growth than younger students did) and the synthetic phonics (OG) group had higher end-of-year scores than the analytic phonics group had (d = .29), but not the sight word group (d = .05). When the other covariates were controlled for, there were no significant differences between the three instructional conditions, and ethnicity was a significant correlate.

A similar pattern was found for word reading. Older students outperformed younger students, and the synthetic phonics (OG) instruction yielded higher word reading than analytic phonics instruction and sight word instruction did (d = .39 and .17, respectively), but those effects were no longer significant when covariates were controlled for; socioeconomic status was the significant correlate in this case. Finally, the effects of the three instructional conditions on word reading were reexamined while controlling for students' initial proficiency with phonological and orthographic processing. There were no longer significant treatment effects once initial skills were included in analysis. However, students who began the school year with higher phonological processing skills developed word reading skills at a higher rate than did students with lower initial skills (significant slope interaction). In sum, without controlling for covariates, the synthetic phonics (OG) instruction led to (a) greater gains in phonological processing and word reading than did both other instructional conditions and (b) greater gains in orthographic processing than did the analytic phonics instruction. However, when demographic variables, verbal IQ, and initial skills were examined as covariates, there were no significant differences between the three instructional methods. These findings suggest that examination of covariates is an important consideration in analyzing the effects of different types of reading instruction and that the effectiveness of one instructional program over another may be influenced by student characteristics.

Joshi, Dahlgren, and Boulware-Gooden (2002) implemented the Language Basics: Elementary curriculum (as cited in Joshi et al., 2002) (based on Cox's 1992 Alphabetic Phonics) in first-grade general education classrooms. Two classes were taught using Language Basics, and their progress was compared to two first-grade classrooms taught using the Houghton Mifflin reading program in a quasi-experimental design. At the end of first grade, both groups demonstrated significant growth in comprehension. Only the OG-based group made significant growth in word attack and phonological awareness. The Language Basics groups had significantly higher scores on word attack (d = 1.08), phonological awareness (d = .44), and comprehension (d = .81).

In the final elementary school–based study, Westrich-Bond (1993) used a quasi-experimental design to examine the effects of OG instruction as compared to basal reading instruction using the Ginn basal series. The participants were students with learning disabilities who received special education services in either resource room classrooms or self-contained special education classrooms. Students in both classroom types received either the OG reading instruction or basal reading instruction for a total of four conditions (resource room with OG reading, resource room with basal reading, self-contained with OG reading, and self-contained with basal reading). Reading instruction occurred during four sessions per week. The *Woodcock Reading Mastery Test* (Woodcock, 1989) Word Identification and Word Attack subtests were used as the measures of real word and non-word reading. There were significant gains scores for word attack and word identification from pretest to posttest (after controlling for initial reading skill), but there were no significant differences between the two instructional conditions (OG vs. basal reading instruction). There were, however, differences in gain scores for the type of special education classrooms. Gains in word attack scores were greater for students in self-contained classrooms than for students in resource room classrooms, while gains in word reading scores were greater for students in self-contained classrooms than for students in resource room classrooms. However, there were no significant differences between the type of reading instruction for real word reading (for resource room d = –.40, and for self-contained special education classroom d = –1.40) or word attack (for resource room d = .07, and for self-contained special education classroom d = .03) after controlling for initial word reading skill. Additional analysis was also conducted that controlled for the number of months of instruction, but there were still no significant difference between the two types of reading instruction (OG vs. basal reading).

Clinic-Based Interventions. Oakland, Black, Stanford, Nussbaum, and Balise (1998) investigated an OG-based remediation program known as the Dyslexia Training Program (DTP), developed at the Texas Scottish Rite Hospital. Based on Alphabetic Phonics (Cox, 1992), DTP offers both teacher-directed instruction and a video-directed instructional program to be used by teachers who have not had extensive training in the program. Using a quasi-experimental design, Oakland et al. compared the use of teacher-directed instruction and video-directed instruction to a control group that participated in the resource room remedial program. Students (mean age = 11 years, SD = 2) received instruction 1 hour per day, 5 days a week for 2 years. Students receiving DTP (regardless of whether it was the teacher-directed or the video-directed program) significantly outperformed the students in the control treatment on measures of comprehension (d = .65), word reading (d = .73), spelling (d = .24) and decoding monosyllabic words (d = .45) and polysyllabic words (d = .80). Students in the DTP group outperformed their counterparts in the control group despite lower initial skills in word reading.

The final elementary-school-age study was conducted by Hook, Macaruso, and Jones (2001), who compared the efficacy of Fast ForWord (a computer-based instructional program designed to improve auditory processing skills) to OG instruction. In this summer program, students (7–12 years old) either received OG reading instruction or participated in Fast ForWord activities. Matched pairs of students (matched by

IQ, phonemic awareness ability, reading level, and age) received one of the intervention treatments. Pre- and posttest assessments included phonemic awareness, word identification, and word attack (additional oral language and phonological processing measures also were collected for the Fast ForWord treatment and a longitudinal control group, but are not discussed here). At the end of the summer program, both groups made significant growth in phonemic awareness (assessed on the *Lindamood Auditory Conceptualization Test*; Lindamood & Lindamood, 1979). For reading, the OG group made significant growth in Word Attack, whereas the Fast ForWord groups did not. Neither group made significant growth in Word Identification. Effect sizes comparing posttest differences indicated small effects for phonemic awareness (d = .11) and word identification (d = .07), but large effects for Word Attack (d = .77).

Middle School and Secondary Students

Two studies investigated the effectiveness of OG with adolescents. Simpson, Swanson, and Kunkel (1992) investigated reading remediation for students (aged 13–18 years old) placed in juvenile detention facilities. In a quasi-experimental design, students who received 90 min per day (5 days per week) of OG-based instruction were compared to a comparison group of students who received 45 min per day of traditional reading instruction in a second detention facility. Of the 116 students who began the study, 63 students were available for posttesting. The OG instruction resulted in more reading growth—as assessed by the *Woodcock Reading Mastery Test* (Woodcock, 1989)—than did the remedial English condition. Because of large differences in the amount of instructional time (both among students and between experimental conditions), additional analysis was conducted to predict the rate of reading growth relative to hours of reading instruction. In this case, students who received OG instruction, on average, outperformed the comparison condition. Analysis of individual cases indicated that not all students made growth. The authors concluded that students were likely to make more growth if they received more hours of instruction. However, given the specific circumstances of the participating students, not all students were at the detention facility long enough to receive an adequate amount of instruction to make growth.

Dooley (1994) compared the effects of reading instruction that incorporated multisensory reading instructional principles and cooperative learning. Using a quasi-experimental design, multisensory integrated reading instruction (MIRC) was compared to traditional instruction for middle school students in remedial reading classes. Students participated in instruction for one semester, and reading and writing outcomes were assessed. For all measures except one (general vocabulary, d = .01), the MIRC group statistically significantly outperformed the control condition (word attack d = .46, reading rate d = .83, syntactic similarities d = .44, paragraph reading d = .42, sentence sequencing d = .51). The intervention in this study also included specific instruction in writing, and students demonstrated statistically significant improvements in writing. Students who participated in MIRC had significantly higher performance on thematic maturity (d = 1.15) and contextual style (d = .79).

College Students

Three studies investigated the effectiveness of OG and OG-based instructional programs for college students. Only one study reported sufficient statistics to calculate effect sizes.

For two studies, the participants were college students enrolled in Marshall University's HELP (Higher Education for Learning Problems) programs. Guyer and Sabatino (1989) reported positive results for a multisensory phonics program for students with dyslexia enrolled in a 5-week summer program. Participants received either phonetic reading instruction using OG instruction, nonphonetic reading instruction using a basic reading skills curriculum, or were part of a non-intervention comparison group. At the end of the summer program, there were significant differences among groups in performance on the Wide Range Achievement Test (WRAT; Jastak & Wilkinson, 1984) Reading subtest and Woodcock Reading Mastery Test (WRMT; Woodcock, 1973) Total Reading score. The authors report that there was differential response to instruction and that the OG groups made greater amounts of progress than did groups with the nonphonetic approach or those in the no-intervention control group. Although there was a relatively small sample size (N = 30), the participants demonstrated gains over a relatively short period of time.

In the second college-level study, Guyer, Banks, and Guyer (1993) investigated the effectiveness of the Wilson Reading System (Wilson, 1996) for improving spelling skills. College students with dyslexia who requested reading and spelling intervention were randomly assigned to two instructional conditions and compared to a non-intervention control. The two intervention groups were provided with a semester (two 1-hour sessions per week for 16 weeks) of spelling instruction using the Wilson program or instruction in a nonphonetic (whole word) approach to spelling. Using the WRAT (Jastak & Wilkinson, 1984) Spelling subtest as the outcome measure, students who received spelling instruction under the Wilson program significantly improved their spelling performance, whereas the nonphonetic approach and no-intervention groups did not demonstrate gains.

In the third college-level study, Chandler, Munday, Tunnell, and Windham (1993) compared OG instruction using Alphabet Phonics (Cox, 1992) to a traditional remedial reading instructional approach, which focused on comprehension skills, reading efficiency, study skills, and test-taking strategies in a quasi-experimental design. Community college students participated in a course for one of two semesters. Chandler et al. reported findings in contrast to those of Guyer et al. (1993), specifically, that the students in the control group, who re-

ceived traditional reading instruction, performed significantly higher than did the students in the OG group. OG instruction did improve the reading performance of college students in developmental reading courses; however, a traditional developmental course using a textbook and focusing on vocabulary and comprehension development was more effective in increasing overall reading levels, as measured by the Nelson Denney Reading Test ($d = -1.18$).

Across the studies that targeted college students, one important finding was that adults with reading disabilities may benefit from continued reading instruction. Guyer et al. (1993) and Guyer and Sabatino (1989) identified OG instruction as more effective than other approaches in improving reading and spelling skills, whereas Chandler et al. (1993) reported that the OG instruction was less effective than a traditional remedial reading course typically implemented by community colleges. All three studies were conducted with a small number of participants, used a limited number of measures to assess effectiveness, and had other methodological flaws that could limit the strength of this evidence.

Summary

The findings of this literature review are summarized in two ways. First, we summarize the results of the effect sizes findings across studies. This allows us to draw conclusions about which reading and reading-related skills are affected by OG and OG-based instructional programs. Second, we summarize the overall outcomes of each study. This provides a more comprehensive picture of how much support exists for the efficacy of OG and OG-based instruction.

Effect Sizes. We coded all studies for the type of reading or reading-related outcome and then classified outcomes into the following categories: word identification (reading real words), word attack or nonword reading (reading phonetically regular nonwords), comprehension, vocabulary, fluency, spelling, total reading, phonological/phonemic awareness, and orthographic awareness. We then calculated the mean effect size for each outcome. (Multiple effects sizes for a single outcome within a single study were averaged prior to being aggregated across studies.)

The greatest effect, on average, for the 12 studies was for word attack or nonword reading, with a mean effect of 0.82, which is considered a large effect (7 studies). Effect sizes ranged from 0.02 to 1.94. For word identification, the mean effect size was 0.42, with effect sizes ranging from −.091 to 1.56 (6 studies). The mean effect size for total reading was 0.37, ranging from −1.18 to 1.59 (4 studies). The mean effect size for comprehension was 0.76, ranging from 0.17 to 1.68 (5 studies). For phonological awareness, the mean effect size was 0.34, ranging from .11 to 0.48 (3 studies). Two studies measured vocabulary ($M\ d = 0.81$, ranging from 0.01 to 1.60), and effect sizes for spelling ($d = 0.24$) could be calculated from 1 study. Effect sizes for fluency could be calculated from 1 study ($d = 0.83$), and orthographic awareness was examined in 1 study (Foorman et al., 1997; $d = .29$ to .05 for the two comparison conditions). The largest mean effect for OG and OG-based reading program was for word attack/nonword reading and comprehension outcomes, with a small to medium effect for real word reading. Other mean effect sizes are in the small to medium range. Mean effect sizes should be interpreted with some caution given the wide range of effect sizes within each outcome category and the use of different measures of each outcome.

Comparison Results. Twelve studies that compared OG or OG-based instructional procedures to another reading approach were reviewed. Five studies (Guyer et al., 1993; Guyer & Sabatino, 1989; Joshi et al., 2002; Litcher & Roberge, 1979; Simpson et al., 1992) reported that OG instruction was more effective than one or both comparison approaches for all outcomes examined in the study. Four additional studies (Dooley, 1994; Hook et al., 2001; Oakland et al., 1998; Stoner, 1991) reported that OG instruction was more effective than one or both of the comparison approaches for at least one, but not all, outcomes. Two studies reported that the alternate instructional condition was more effective (Chandler et al., 1993; Westrich-Bond, 1993). Finally, Foorman et al. (1997) reported that relative effectiveness of OG instruction depended on the examination of covariates, and initial findings of the superiority of OG instruction to analytic phonics instruction and/or sight word reading instruction was reduced when covariates were controlled for.

Discussion

The purpose of this paper was to review the empirical research on OG and OG-based reading interventions. Twelve studies (10 published articles and 2 dissertations) provided what could be best described as a preliminary evidence base of the effectiveness of OG and OG-based programs in comparison to other instructional approaches. A synthesis of the effectiveness of OG and OG-based reading instruction, the need for methodological rigor in the study of such programs, and implications for future research and practice are discussed in the following sections.

Effectiveness of OG and OG-Based Programs

The central purpose of this literature review was to summarize the scientific research to determine the effectiveness of OG and OG-based instructional programs. Studies that cited positive outcomes for OG and OG-based instruction reported effects on various outcomes. Positive results were found for word reading (Guyer & Sabatino, 1989; Hook et al., 2001; Litcher & Roberge, 1979; Stoner, 1991), word attack/decoding (Hook et al., 2001; Joshi et al., 2002; Litcher & Roberge, 1979;

Stoner, 1991), spelling (Guyer et al., 1993; Oakland et al., 1998), and comprehension (Joshi et al., 2002; Litcher & Roberge, 1979; Oakland et al., 1998; Stoner, 1991). Positive outcomes for OG and OG-based instruction were reported across settings and populations. For beginning readers, positive effects were reported for first-grade children in general education classrooms, for elementary children at risk for or identified with reading disabilities and served in public schools, and elementary-age children in clinical settings. Positive effects were also found in two studies of college students with diagnosed learning disabilities (Guyer et al., 1993; Guyer & Sabatino, 1989).The findings were not, however, all positive in favor of OG instructional programs. Nor were the findings statistically significant favoring either instructional program. For example, Chandler et al. (1993) reported that the alternate instructional condition was more effective than OG instruction. In addition, there were nonsignificant results for all participants reported by Westrich-Bond (1993), for participants in second and third grade reported by Stoner (1991), and for specific outcomes reported by Oakland et al. (1998) and by Hook et al. (2001).

We find it interesting that effects on vocabulary were reported in only two studies and that only one study examined reading fluency as an outcome of instruction. Since being identified by the National Reading Panel (2000) report as two of five essential components for reading, these two areas are currently receiving increased research attention. It is important to note that the primary components of OG instruction typically do not explicitly address these areas of instruction, although several programs do include opportunities for instruction or practice and/or suggestions in these areas. For OG programs to demonstrate their effectiveness for all aspects of reading, perhaps they will need to be expanded or supplemented with other reading instruction approaches that either build vocabulary and fluency or more systematically study these specific outcomes.

Despite the wide acceptance and enthusiasm for OG and OG-based programs, not all studies reported them to be superior, and caution should be taken when attempting to generalize any of the reviewed results. Differences in study participants, setting, location, program type, instruction time, the OG instructional programs and implementation, and outcome measures must be considered when evaluating this research.

Need for Rigorous Scientifically Based Research of OG

The process of reviewing and summarizing OG and OG-based reading instruction research was challenging due to the small number of extant studies that employed experimental or quasi-experimental designs. Several studies were eliminated from this review because only pretest and posttest data were reported, because the report was described as a pilot study, or because the study included a small number of participants (i.e., $n < 10$). Given the popularity of OG and OG-based instructional methodologies, we were surprised that only a dozen studies met the inclusion criteria of our review. We used a rather liberal inclusion criterion to allow for examination of the majority of extant studies. If we were to require a more stringent criterion, the scant number of studies would provide little information from which to draw any conclusions.

Additionally, methodological issues present in many of the reviewed studies complicate any conclusions that could be drawn. All but 1 of the 12 studies used a quasi-experimental design, in which groups or classrooms of students (instead of individual students) were assigned to experimental conditions. Seven out of 12 studies had small sample sizes (total samples $n < 50$, see Table 1). Several studies were published in the 1970s and 1980s, when standards for educational research were less stringent and when technology for statistical analysis was not readily available to researchers. Reports of these older studies did not include components that are considered compulsory in any high-quality research report today (Gersten et al., 2005). Several of the more recent articles also did not report such information as details regarding the procedures used to ensure comparability or equivalency of treatment groups in quasi-experimental designs, fidelity with which instruction was implemented, technical characteristics of the measures used as dependent variables, or scope of training provided to instructors. The absence of such information makes it impossible to determine the internal validity of the research and suggests that, at this time, there is insufficient evidence to conclude that OG and OG-based reading instruction meet the requirements of scientifically-based reading instruction. The methodological rigor of future OG research must be improved or strong conclusions regarding its effectiveness cannot be drawn.

This review of literature also raises areas for future research. The majority of the studies simply identify whether OG instruction is more effective than the comparison treatment(s). Given the number of variations in OG and OG-based programs, research examining the relative effectiveness of these programs seems warranted. Research that examines instructional setting (e.g., individual, small-group, whole class instruction), as well as the effectiveness of different components of OG (e.g., phonology and phonological awareness, sound-symbol correspondence, and syllabication), would help to specify the conditions under which OG instruction is effective. Finally, it is important to identify for whom OG instruction is most effective and the types of responses that can be expected for students with different instructional needs.

Limitations

A limitation of the current review was the exclusion of research that was not published in peer-refereed journals or a doctoral dissertations. Conference presentations, book chapters, and research available through other venues (i.e., Web sites, publisher information, and other reports) were not included, but may provide additional support for the efficacy of

OG and OG-based reading instruction. However, peer review and publication in refereed journals is a necessary condition to determine the methodological rigor of research studies. The studies reviewed here have met at least the minimum requirements of the journal in which they were published or had been conducted with the expert consultation of a dissertation committee.

Summary and Conclusions

In sum, the extant research literature provides both evidence that supports, as well as evidence that fails to support, the effectiveness of OG instruction in reading, when compared to other reading instruction. This review also highlights the disparity between research and practice. For decades, educational researchers have lamented the "research to practice gap"—that is, the reluctance of classroom teachers to implement educational practices that have been developed and validated by researchers (see. Gersten, Chard, & Baker, 2000; Greenwood & Abbott, 2001). In the case of OG instruction, the opposite phenomenon—a "practice to research gap"—appears to exist. Despite widespread use by teachers in a variety of settings for more than 5 decades, OG instruction has yet to be comprehensively studied and reported in peer-refereed journals. The small number of existing studies lack methodological rigor that would be required for publication in current peer-refereed journals.

The major conclusion of the review is that the research is currently inadequate, both in number of studies and in the quality of the research methodology, to support that OG interventions are scientifically based. Given the inconclusive and mixed nature of the extant research findings, it may be premature to reconsider the implementation and use of OG reading instruction programs for children with reading disabilities. OG instructional programs continue to be implemented by teachers who find it an effective reading instruction program in a variety of settings. However, it is necessary to scientifically determine (a) if OG and OG-based instructional interventions are effective for all students with reading disabilities, (b) for which children OG reading instruction is most effective, (c) the effectiveness of OG-based reading programs in comparison to other remedial approaches, and (d) the specific conditions under which OG is most effective. Only then can the scientific basis of such reading instruction be assured.

REFERENCES

Bhat, P., Rapport, M. J., & Griffin, C. C. (2000). A legal perspective on the use of specific reading methods for students with learning disabilities. *Learning Disability Quarterly, 23,* 283–297.

Biasotto, V. L. (1993). Project ASSIST Institute: An Orton-Gillingham/Spalding based curriculum for teachers and volunteers. *Annals of Dyslexia, 43,* 260–270.

Brown, V., Hammill, D. D., & Wiederholt, J. (1978). *The Test of Reading Comprehension*. Austin, TX: PRO-ED.

Chandler, C. T., Munday, R., Tunnell, J. W., & Windham, R. (1993). Orton-Gillingham: A reading strategy revisited. *Reading Improvement, 30,* 59–64.

Clark, D. B., & Uhry, J. K. (1995). *Dyslexia: Theory and practice of remedial instruction*. Baltimore: York Press.

Cohen, J. (1988). *Statistical power for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cox, A. R. (1992). *Foundations for literacy: Structures and techniques for multisensory teaching of basic written language skills*. Cambridge, MA: Educators Publishers Service.

Dooley, B. (1994). *Multisensorially integrating reading and composition: Effects on achievement of remedial readers in middle school*. Unpublished doctoral dissertation, Denton: Texas Woman's University..

Edmark Reading Program. (1984). Redwood, WA: Edmark Corporation.

Enfield, M. L. (1971). Specific language disability programs in Bloomington, Minnesota Public Schools. *Bulletin of the Orton Society, 21,* 95–110.

Enfield, M. L., & Greene, V. (1997). *Project read.* Bloomington, MN: Language Circle Enterprise.

Fast ForWord. (1998). Berkely, CA: Scientific Learning Corporation.

Foorman, B. R., Francis, D. J., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J. M. (1997). Early interventions for children with learning disabilities. *Scientific Studies in Reading, 1,* 255–276.

Gersten, R., Chard, D., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33,* 445–457.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators of group experimental and quasi-experimental research in special education. *Exceptional Children, 71,* 149–164.

Gillingham, A., & Stillman, B. W. (1960). *Remedial training for children with specific disability in reading, spelling, and penmanship*. Cambridge, MA: Educators Publishing Service.

Gillingham, A., & Stillman, B. W. (1997). *The Gillingham manual: Remedial training for children with specific disability in reading, spelling, and penmanship*. (8th ed.). Cambridge, MA: Educators Publishing Service.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education, 24,* 276–289.

Guyer, B. P., Banks, S. R., & Guyer, K. E. (1993). Spelling improvement for college students who are dyslexic. *Annals of Dyslexia, 43,* 186–193.

Guyer, B. P., & Sabatino, D. (1989). The effectiveness of a multisensory Alphabetic Phonics approach with college students who are learning disabled. *Journal of Learning Disabilities, 22,* 430–434.

Hammill, D., & Larsen, S. C. (1988). *The test of written language–2nd edition*. Austin, TX: ProEd.

Herman, R. (1993). *The Herman method for reversing reading failure*. Sherman Oaks, CA: Herman Method Institute.

Hook, P., Macaruso, P., & Jones, S. (2001). Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties: A longitudinal study. *Annals of Dyslexia, 51,* 75–96.

Hutcheson, L., Selig, H., & Young, N. (1990). A success story: A large urban district offers a working model for implementing multisensory teaching into the resource and regular classroom. *Annals of Dyslexia, 40,* 79–96.

Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test-Revised*. Wilmington, DE: Jastak Associates.

Joshi, R. M., Dahlgren, M., & Boulware-Gooden, R. (2002). Teaching reading in an inner city school through a multisensory approach. *Annals of Dyslexia, 52,* 229–242.

Karlsen, B. R., & Gardner, E. (1985). *Stanford diagnostic reading test*. San Antonio, TX: Psychological Corporation.

Knight, J. K. (1995). Starting over: An overview. In C. W. McIntyre & J. S. Pickering (Eds.), *Clinical studies of multisensory structured language education* (pp. 231–232). Salem, OR: International Multisensory Structured Language Education Council.

Lindamood, C., & Lindamood, P. (1979). *The Lindamood auditory conceptualization test*. Austin, TX: PRO-ED.

Litcher, J. H., & Roberge, L. P. (1979). First grade intervention for reading achievement of high risk children. *Bulletin of the Orton Society, 24*, 238–244.

MacGinitie, W. H., & MacGinitie, R. K. (1989). *Gates-MacGinitie reading tests*. Itasca, IL: Riverside.

McIntyre, C. W., & Pickering, J. S. (Eds.). (1995). *Clinical studies of multisensory structured language education*. Salem, OR: International Multisensory Structured Language Education Council.

No Child Left Behind. (2002). Retrieved September 27, 2004, from http://www.ed.gov/policy/elsec/leg/esea02.html

Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An evaluation of the Dyslexia Training Program: A multisensory method for promoting reading in students with reading disabilities. *Journal of Learning Disabilities, 31,* 140–147.

Ogden, S., Hindman, S., & Turner, S. D. (1989). Multisensory programs in the public schools: A brighter future for LD children. *Annals of Dyslexia, 39,* 247–267.

Orton, S. (1937). *Reading, writing, and speech problems in children*. New York: Norton.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Thousand Oaks, CA: Sage.

Simpson, S. B., Swanson, J. M., & Kunkel, K. (1992). The impact of an intensive multisensory reading program on a population of learning disabled delinquents. *Annals of Dyslexia, 42,* 54–66.

Slingerland, B. H., & Aho, M. (1994–1996). *Slingerland reading program*. Cambridge, MA: Educators Publishing Service.

Spalding, R. B., & Spalding, W. T. (1990). *The writing road to reading* (4th ed.). New York: Quill.

Sparks, R., & Ganschow, L. (1993). The effects of multisensory structured language instruction on native language and foreign language aptitude skills of at-risk high school foreign language learners: A replication and follow-up study. *Annals of Dyslexia, 43,* 194–216.

Sparks, R., Ganschow, L., Kenneweg, S., & Miller, K. (1991). Use of an Orton-Gillingham approach to a foreign language to dyslexic/learning-disabled students: Explicit teaching of phonology in a second language. *Annals of Dyslexia, 41,* 96–118.

Sparks, R., Ganschow, L., Pohlman, J., Skinner, S., & Artzer, M. (1992). The effects of multisensory structured language instruction on native language and foreign language aptitude skills of at-risk foreign language learners. *Annals of Dyslexia, 42,* 25–53.

Stoner, J. (1991). Teaching at-risk students to read using specialized techniques in the regular classroom. *Reading and Writing. An Interdisciplinary Journal, 3,* 19–30.

Torgesen, J. K., & Bryant, B. R. (1994). *The test of phonological awareness*. Austin, TX: PRO-ED.

Traub, N., & Bloom, F. (1992). *Recipe for reading* (3rd ed.). Cambridge, MA: Educators Publishing Service.

Vickery, K. S., Reynolds, V. A., & Cochran, S. W. (1987). Multisensory teaching approach for reading, spelling, and handwriting: Orton-Gillingham based curriculum in a public school setting. *Annals of Dyslexia, 37,* 189–200.

Westrich-Bond, A. (1993). *The effect of direct instruction of a synthetic sequential phonics program on the decoding abilities of elementary school learning disabled students*. Unpublished doctoral dissertation, New Brunswick: The State University of New Jersey, Rutgers.

Wilson, B. A. (1996). *Wilson reading system*. Millbury, MA: Wilson Language Training Corporation.

Woodcock, R. W. (1973). *Woodcock reading mastery test*. Circle Pines, MN: American Guidance Services.

Woodcock, R. W. (1989). *Woodcock reading mastery test–Revised*. Circle Pines, MN: American Guidance Services.

Woodcock, R. W. (1998). *Woodcock reading mastery test–Revised/Normative Update*. Circle Pines, MN: American Guidance Services.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson psychoeducational battery–Revised*. Chicago: Riverside.