

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/249835288>

# Improving the Reading Skills of Urban Elementary Students Through Total Class Peer Tutoring

Article in Remedial and Special Education · April 2007

DOI: 10.1177/07419325070280020801

CITATIONS

24

READS

176

3 authors:



[Lefki Kourea](#)

University of North Carolina at Charlotte

8 PUBLICATIONS 109 CITATIONS

[SEE PROFILE](#)



[Gwendolyn Cartledge](#)

The Ohio State University

52 PUBLICATIONS 686 CITATIONS

[SEE PROFILE](#)



[Shobana Musti-Rao](#)

Pace University, Pleasantville, NY, United States

27 PUBLICATIONS 172 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



TaSDi-PBS [Erasmus Plus: 2016-1-HR01-KA201-022147] [View project](#)



Special issue [View project](#)

All content following this page was uploaded by [Lefki Kourea](#) on 19 July 2014.

The user has requested enhancement of the downloaded file. All in-text references [underlined in blue](#) are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

## Improving the Reading Skills of Urban Elementary Students Through Total Class Peer Tutoring

LEFKI KOUREA, GWENDOLYN CARTLEDGE, AND SHOBANA MUSTI-RAO

### ABSTRACT

This study reports the results of a peer-mediated intervention, total class peer tutoring, on the academic performance of six urban students at risk for reading failure. A multiple baseline design across subjects was used to evaluate the effects of this intervention. The results showed that five of the six students significantly increased their sight-word acquisition and maintenance. All target students' reading fluency and comprehension scores on the standardized *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) Oral Reading Fluency (DORF) passages were higher during intervention than at baseline. Greater fluency gains were shown on constructed paragraphs that included the tutoring sight words. More modest fluency and comprehension gains were found in DORF passages that did not contain words taught in the intervention. Social validity measures taken from teachers, parents, and students revealed positive evaluations of the intervention. Limitations, future directions in research, and practical implications are discussed.

STUDENTS IN URBAN SETTINGS ARE OFTEN AT A disadvantage educationally compared to their peers in suburban areas (Cartledge, 2002). Urban learners are exposed to a variety of family, neighborhood, school, and societal risk factors such as poverty, abuse, neglect, and ineffective academic instruction, which lead to academic problems and contribute to maladaptive behaviors (Gottlieb, Alter, Gottlieb, & Wish-

ner, 1994). Consequently, these behavioral and learning problems result in negative outcomes such as low academic achievement, increased school discipline referrals, and overrepresentation in classrooms serving students with learning disabilities, mental retardation, and behavior disorders (Gottlieb et al., 1994). Furthermore, the aforementioned problems are associated with long-term poor postschool outcomes. Urban minority students are more likely to experience higher rates of school failure, delinquency, lower employment rates, and overall lower adjustment in adulthood than their European American peers without disabilities (Ferri & Connor, 2005).

A number of best teaching practices have emerged in order to prevent urban minority school children from further failure. Best practice refers to the use of research-based instructional strategies and organizational procedures that demonstrate positive outcomes on students' learning (L. J. Miller & Kohler, 1993). One promising instructional program is *total class peer tutoring* (Lo & Cartledge, 2004), also referred to as *classwide peer tutoring*, which accounts and compensates for family and sociocultural risk factors and prevents students' early academic failure (Greenwood & Delquadri, 1995).

Total class peer tutoring is a specific form of peer-mediated intervention in which students are provided with ample opportunities for active engagement and practice. As a

result, students increase their on-task behavior, receive individualized instruction and immediate error correction with positive feedback, and, more important, improve their academic and social skills (Arreaga-Mayer, 1998; A. D. Miller, Barbetta, & Heron, 1994). It is an effective and excellent teaching tool for the inclusion of students with disabilities in general education classrooms. In essence, total class peer tutoring enhances and supports the learning of *all* students with and without disabilities in mainstream settings.

Research studies have well-documented the positive effects of total class peer tutoring intervention on the performance of low-achieving students' sight-word vocabulary (Butler, 1999; Heron, Heward, Cooke, & Hill, 1983), spelling (Delquadri, Greenwood, Stretton, & Hall, 1983; Maheady & Harper, 1987), reading fluency (Kamps, Barbetta, Leonard, & Delquadri, 1994; Simmons, Fuchs, Fuchs, Hodge, & Mathes, 1994), Spanish vocabulary (Wright, Cavanaugh, Sainato, & Heward, 1995), social studies (Lo & Cartledge, 2004), and functional math skills (Arreaga-Mayer, 1998). Along the same lines, similar academic and social benefits have been demonstrated for students with moderate to severe disabilities (McDonnell, Mathot-Buckner, Thorson, & Fister, 2001), English language learners (Greenwood, Arreaga-Mayer, Utley, Gavin, & Terry, 2001), students with learning disabilities (Simmons et al., 1994), attention-deficit disorder (DuPaul & Henningson, 1993), mental retardation (Heron et al., 1983), autism (Kamps et al., 1994), and behavioral disorders (Lazerson, 1980). The superior effects of peer tutoring have been extended to a variety of populations, including kindergarten (Brady, 1997), elementary (Heward, Heron, & Cooke, 1982), middle school (Nazzari, 2002), high school (Maheady, Sacca, & Harper, 1987), and college students (Fantuzzo, Riggio, Connelly, & Dimeff, 1989).

Although total class peer tutoring has been identified as a successful intervention that addresses students' academic deficits, many research questions remain unanswered with respect to its effects on sight-word acquisition and, consequently, on oral reading fluency. Some studies have investigated reading fluency as an instructional component of reciprocal peer tutoring and its effects on students' reading achievement. Simmons et al. (1994) showed that students increased their fluency and comprehension after participating in classwide reciprocal peer tutoring, in which students read basal texts and their tutors corrected word recognition errors by following a corrective procedure. Likewise, Kamps et al. (1994) found that students increased their reading rate and comprehension after reading passages when they received feedback and reinforcement from peers. Students' fluency as a measure of generalizing tutored sight words into context has yet to be examined in the literature. In other words, having students practice a number of sight words in reciprocal peer tutoring and asking them to generalize (i.e., read) these words in untaught passages deserves careful examination. The purpose of sight-word acquisition is to enable students to build a basic vocabulary and use it in a reading context. To

the authors' knowledge, no previous studies have been conducted investigating how the increase of sight-word learning affects students' reading in context.

Furthermore, although there is ample empirical support for the effects of total class peer tutoring (e.g., Heron et al., 1983) on students' maintenance of sight words, additional research needs to focus on the time period for which students can maintain tutored words. Heron et al. (1983) found that first-grade students could maintain a mean of 89% of tutored words 1 week after sight words had been mastered. Given the critical importance of word knowledge in students' reading performance, maintenance of sight words more than one week after intervention needs to be examined more closely.

The current study was designed to extend previous research on total class peer tutoring by investigating the effects of this instructional strategy on students' sight-word acquisition. Moreover, this study sought to examine if students were able to maintain words for 2 to 3 weeks and 17 to 20 weeks after first being introduced to the words. Researchers also examined students' reading fluency and comprehension as a generalization measure of sight-word recognition in untaught passages.

## METHOD

### Setting and Participants

The study was conducted in an inclusive second/third-grade (i.e., joint) classroom of an urban elementary school located in a midwestern metropolitan area. The school had an enrollment of 148 students from preschool through fifth grade. The majority of the school population consisted of African American students (84.5%). The remainder of the population was composed of 12.8% European Americans, 1.3% Asian Americans, 0.7% Native Americans, and 0.7% Hispanic Americans. The targeted classroom included 14 African American students (8 girls and 6 boys). More than half of the class (i.e., 8 students) was receiving special education services outside the general education classroom 50% to 60% of the school day.

During the intervention, both the general and the special education teachers were present and helped to facilitate the peer tutoring. When peer tutoring was applied at the class-wide level, the general education teacher was the main implementer of the program, whereas the special education teacher was monitoring students during the intervention. The general education teacher was informed about the basic procedures of peer-mediated interventions because she had been involved in similar peer-mediated interventions (e.g., paired repeated readings, reciprocal peer tutoring) the previous year. Nevertheless, a separate meeting was held between the first author and the general education teacher prior to the class-wide intervention to discuss the student training procedures for peer tutoring. Both the special education and general education teachers were co-teaching in the inclusive urban

classroom for 2 to 3 hours each day. However, special education and general education students were separated during the primary subject courses (e.g., math, language arts) and were taught by their respective teachers. The special education teacher had more than 15 years of teaching experience, whereas the general education teacher had only 3 years of teaching experience. Peer tutoring sessions were held three times per week for 30 min. They were conducted on Mondays and Tuesdays from 10:00 to 10:30 a.m. and on Wednesdays from 1:30 to 2:00 p.m.

Six out of 14 African American students (5 second graders and 1 third grader) were identified for this study. Target students ages 7 to 8 years, with a mean age of 7 years 4 months, were selected at the beginning of the study. Two students received special education services under the learning disabilities category, whereas the other 4 general education students were identified as at risk. The 6 students were included in this study because of (a) low performance on the four standardized subtests (Letter-Word Identification, Reading Fluency, Passage Comprehension, and Word Attack) of the *Woodcock-Johnson-III Tests of Achievement* (WJ-III; Woodcock, McGrew, & Mather, 2001) and (b) participation in interventions during the previous years of the project. The focus of the project had been to increase the academic performance of urban minority students through culturally responsive interventions. The current intervention was an effort to provide follow-through with the same students. Parental consent was obtained for all students who participated in the study; however, data were collected only for the six target students. A description of the demographic characteristics of the target students is presented in Table 1.

## Dependent Variables

The study focused on four dependent variables: sight-word acquisition, reading fluency, comprehension, and maintenance.

**Sight-Word Acquisition.** A sight word was considered acquired when the word was read correctly by the student within 3 seconds after the word had been presented on a plain 3-inch  $\times$  5-inch index card. Incorrect responses were recorded if the student uttered a different word, responded after 3 seconds, or made no response. Variations in the student's pronunciation, articulation, and dialect were not counted as errors. For example, if student said "thoouth" when the word *those* was presented, then the response was still counted as correct. Mastery of sight words was considered achieved at the last session of the week. Students received a different set of sight words for practice at the beginning of each week.

**Reading Fluency.** Students' reading fluency rate was defined as the number of words correctly read per minute in a passage. The first author, who served as the primary experimenter in this study, used two sets of passages. The first set

consisted of standardized grade-level passages from the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) Oral Reading Fluency (DORF; Good & Kaminski, 2002) measure. Second-grade students were assessed on the second-grade-level DORF passages, and the third grader was assessed on the third-grade-level DORF passages. The second set of passages, known as *constructed paragraphs*, was developed by the experimenters and included sight words that students had practiced during intervention.

To evaluate reading fluency, the experimenter followed the scoring rules defined by the authors of the DORF passages (Good & Kaminski, 2002) for both sets of passages. Specifically, a word was considered correct if the student (a) read the word correctly in the context of the sentence within 3 seconds of the word's presentation and (b) corrected the word within 3 seconds following an error. A word was defined as incorrect when it was (a) omitted, (b) read correctly but in the wrong order (e.g., the sentence was "I drank too much" and the student read "I too drank much." The words *drank* and *too* were counted as incorrect, although they were read correctly, because the word order was wrong), (c) mispronounced in the sentence, and (d) read or not for more than 3 seconds (e.g., student hesitated or struggled with words). Repetitions, insertions, and words of imperfect pronunciation because of dialect, articulation, or second language interference were not counted as either correct or incorrect, following Good and Kaminski (2002). Furthermore, a discontinuation rule was followed, as outlined in the DORF scoring guide: If the student did not read any words correctly during the first row of the passage, the experimenter discontinued the passage and recorded a score of 0.

**Reading Comprehension.** Student reading comprehension was measured by the number of missing words that the student could identify correctly in each DORF passage

TABLE 1. Target Student Demographic and Academic Characteristics

Student	Grade level	Gender	Age <sup>a</sup>	Disability
Erin	2	Female	7-6	LD
Steve	2	Male	7-1	At risk
Irena	2	Female	7-5	At risk
Dignity	2	Female	7-5	At risk
Dan	3	Male	8-3	At risk
Susan	2	Female	7-1	LD, ADHD

Note. All target students were African American and had low socioeconomic status. LD = learning disabilities; ADHD = attention-deficit/hyperactivity disorder.

<sup>a</sup>Expressed as years-months.

and constructed paragraph. Reading comprehension was measured using the cloze procedure (Grant, 1979). The cloze procedure was selected for three reasons: (a) students had to simultaneously process semantic (word meaning) and syntactic (word order) clues for completing the cloze passages, (b) the cloze procedure provided a quick estimate of the relative difficulty of a particular text for students, and (c) the cloze procedure also was used in the pretest and posttest standardized Passage Comprehension subtest of the WJ-III. Therefore, the experimenter kept the comprehension assessment procedure consistent throughout the study.

Five words were identified for comprehension in each passage and were deleted from the sentences. Three criteria were used to identify and delete sight words from the DORF passages and the constructed paragraphs:

1. Words had to present grammatical functions (e.g., nouns, indefinite articles, verbs, conjunctions). Adverbs or adjectives were avoided in the comprehension assessment because they included the possibility of having more than one correct response for each blank space.
2. Words were not so likely to be substituted with synonyms or would have a very narrow range of possible answers for each missing word. For example, the word *to* was deleted from the sentence "My older sister had \_\_\_\_ go to work," as opposed to the word *older*, as in "My \_\_\_\_ sister had to go to work."
3. Words had to present meaning in the sentence.

A comprehension item was recorded as correct if the student identified (a) the exact original missing word within 5 seconds, either from the first or the second count of reading the sentence, (b) a word that was similar to the original word (i.e., synonyms), and (c) any word of the correct response class that made semantic sense. For instance, in the sentence "My best friend \_\_\_\_ Tim," either *was* or *is* was counted as correct. An incorrect response was defined as one in which the student (a) identified a word that did not match in the context of the sentence or (b) did not provide a word within 5 seconds of reading the sentence for the second time. For example, if the sentence was "My best friend \_\_\_\_ Tim," and the student identified *the* as the missing word, then the answer was scored as incorrect. If the student responded with the word *is* more than 5 seconds after being presented with the sentence for the second time, the response was scored as incorrect. If the student said more than one word, the experimenter prompted the student to say only one word. If the student identified a similar word, the experimenter requested the student to try another word. In the previous example "My best friend \_\_\_\_ Tim," if the student gave the answer *was* as opposed to the original word *is*, the experimenter would tell the student, "The answer could be *was*. Try another word." If

the student did not respond or gave a word different from the original answer (i.e., *is*) on the second try, the answer was considered incorrect.

**Maintenance.** Correct reading of 10 sight words taught during peer tutoring was assessed 2 to 3 weeks after the intervention. Each assessment session consisted of 20 sight words, of which 10 words were presented 2 weeks after peer tutoring had elapsed and another set of 10 words was presented after 3 weeks (except for the initial three assessment sessions, where words were presented after 6, 5, and 4 weeks of peer tutoring, respectively). A sight word was defined as maintained if the student read the word correctly within 3 seconds of presentation of the word printed in black ink on a 3-inch × 5-inch index card. The words that were counted as incorrect were reintroduced to the student during the subsequent weekly pretests and were also used in the set of tutoring words for that week. Furthermore, a cumulative end-of-study assessment was conducted to verify students' retention.

## Instrumentation

**Woodcock-Johnson-III Tests of Achievement.** Four WJ-III subtests were used:

- The Letter-Word Identification subtest measured students' word identification skills.
- The Reading Fluency subtest measured students' ability to quickly read simple sentences, to decide if they are true or false, and to circle the right answer.
- The Passage Comprehension subtest initially evaluated the student's ability to match a picture of an object with a rebus. These items were in a multiple choice format that required students to point to the picture represented by a phrase. The remaining items required students to read a short passage and identify a missing keyword.
- The Word Attack subtest measured students' skill in applying phonic and structural analysis skills for reading aloud unfamiliar printed words.

Woodcock et al. (2001) reported a median reliability in the 5- to 19-year age range of .91 for the Letter-Word Identification subtest, .90 for the Reading Fluency subtest, .83 for the Passage Comprehension subtest, and .87 for the Word Attack subtest.

**DORF Progress Monitoring Passages.** The DIBELS Oral Reading Fluency progress monitoring passages were used to assess students' progress on reading fluency. Test-retest reliability ranged from .92 to .97 for elementary stu-

dents. Criterion-related validity ranged from .52 to .91 (Good & Kaminski, 2002).

**Constructed Paragraphs.** Five passages were developed by the first author, and each passage included two paragraphs. Each of the 10 paragraphs was administered separately. All reading passages contained 20 basic sight words that students had mastered during the intervention. For the first three passages, all target students had 100% of the 20 sight words in their individual bank of peer tutoring words. For Passages 4 and 5, all students but one (i.e., Irena) had at least 50% of these 20 sight words in their bank of words. Irena had 9 out of 20 sight words (40%) included in the paragraphs. Nevertheless, having fewer mastered peer tutoring words did not influence Irena's fluency rate on these paragraphs. On the contrary, Irena was the only second grader who mastered advanced grade-level sight words during intervention. Therefore, it was difficult for the experimenter to include such words (i.e., fourth- to fifth-grade-level sight words) in the constructed paragraphs, because none of the other second-grade target students had advanced to those levels. Paragraphs were tested for readability based on the Flesch-Kincaid grade level that accompanied the Microsoft Word toolbar. The readability grade level of the constructed paragraphs ranged from 2.3 to 3.7.

### Interobserver Agreement

The percentage of interobserver agreement (IOA) for the four dependent variables was calculated by dividing the number of agreements between the two observers (i.e., the experimenter and a second observer) with the total number of agreements and disagreements multiplied by 100. The mean IOA for each dependent variable was calculated by summing all the percentages of IOA across sessions and then dividing the sum by the number of sessions in which the second observer had been present. A mean IOA of 99.1% (range = 92%–100%) was obtained for sight-word acquisition for 36% (25 out of 69) of the study sessions. Two mean agreements were calculated for Reading Fluency, because two types of instruments had been incorporated in the study. For the DORF passages, the mean IOA on Reading Fluency was 96.2% (range = 91%–100%), and for the constructed paragraphs, the mean IOA was 96.7% (range = 95%–97.5%), measured on at least 40% of the sessions. Likewise, there were two IOA measures for reading comprehension. For the DORF passages, the mean IOA was 99.6% (range = 96%–100%), and for the constructed paragraphs, the mean IOA was 99% (range = 96%–100%), measured on at least 40% of sessions. Finally, the mean IOA for maintenance was 98.7% (range = 98%–99%) for 44% of the sessions.

Weekly pretests were conducted across experimental conditions to identify students' unknown sight words. An unknown word was defined as a word that the student was not able to read correctly within 3 seconds after the word had

been presented visually. The mean agreement of weekly pretests was 98.5% (range = 95%–100%) for 35% of the study sessions.

### Treatment Integrity

Procedural integrity checklists were used to evaluate (a) the accuracy of tutors' behavior during peer tutoring and (b) the experimenter's behavior for accurately implementing the steps of reading fluency and comprehension assessments during baseline and intervention. The integrity assessment for reading fluency was obtained from the official administration manual of DIBELS (Good & Kaminski, 2002), and the integrity assessment for reading comprehension was developed by the experimenter. The treatment integrity on tutor and experimenter behaviors was measured by dividing the number of behaviors performed correctly by the total number of required behaviors multiplied by 100.

**Tutor Behavior.** Target students were observed for all tutoring sessions, but they were selected for procedural integrity checks for at least 30% (32 out of 60) of these sessions. All students but one were chosen for fidelity checks at least four times during the study. Susan was observed only three times. However, during these three observations, her mean accuracy was 100%. During intervention, a group mean of 96.7% (range = 92.7%–100%) for treatment integrity was obtained for 32 fidelity checks. Erin and Irena had a mean accuracy of 92.7% and 93.7%, respectively. Erin failed to prompt her partner, and Irena did not praise her partner during practice time. After procedural integrity checks had been completed, during the next tutoring session, the experimenter prompted both students to follow all steps of tutor behavior.

**Experimenter Behavior.** Reading fluency and comprehension procedural integrity checks were conducted for eight sessions. A mean accuracy of 99% (range = 93%–100%) of the experimenter behavior was obtained for eight fidelity checks. In one of the checks, the second observer recorded that the experimenter did not completely perform the standardized verbal directions for reading fluency.

### Observer Training

To reduce the likelihood of having any extraneous variables, such as observer bias or assessment complexity, influencing the IOA measurements (Kazdin, 1977), observer training sessions were conducted for two second observers (undergraduate and graduate students). During training sessions, explicit definitions of the target behaviors were presented, and a precise and clear description of the peer tutoring intervention was provided. Examples of reading passages were shown, and the recording procedure with all its notations was demonstrated. Examples and nonexamples of reading comprehension answers were also given. The procedural integrity

checklist for reading fluency and reading comprehension was given and discussed with second observers in order to avoid any observation drift from the procedures being measured.

## Experimental Design and Conditions

A multiple-baseline across-subjects design was used to evaluate the effects of total class peer tutoring on student performance (Baer, Wolf, & Risley, 1968). According to Cooper, et al. (1987), a multiple baseline design enables the concurrent measurement of the target behavior across multiple subjects and allows direct monitoring for generalization of behavior change. Furthermore, this design does not require the withdrawal of an effective intervention to demonstrate experimental control.

The six target students were paired in the intervention based on similar performance on the initial Letter-Word Identification subtest and the daily sight-word assessments during the baseline condition. Therefore, the pair of students that demonstrated the lowest stable levels of responding during the teacher-led instruction condition first entered the intervention phase. The first pair (i.e., Erin and Steve) remained in baseline for 3 weeks. As soon as the first pair increased the number of sight words learned during peer tutoring, the second and third student pairs entered intervention—after 2 and 3 weeks, respectively. Student pairs remained in intervention as follows: 20 weeks for the first pair, 18 weeks for the second pair, and 17 weeks for the third pair. When all target students had entered intervention, changes were made with the second and third pairs.

**Pretest.** At the beginning of the study, students were tested on the four WJ-III subtests: Letter-Word Identification, Reading Fluency, Passage Comprehension, and Word Attack.

**Baseline.** The experimenter conducted weekly pretests to determine the number of sight words that students knew prior to the teacher's instruction of the five sight words. During each week, the teacher presented five sight words from her predetermined sight-word vocabulary list. Teacher-led instruction consisted of presenting (i.e., saying) sight words in small groups and then spelling the words. Later, the teacher would ask students to spell each word in unison and to complete worksheets using the words.

Data were collected on students' performance three times per week. At the end of the teacher's instruction, students were pulled out individually and were asked to identify the words presented on the flashcards. In addition to sight-word assessment, students were tested on reading fluency and comprehension. For this purpose, the grade-level DORF passages were administered in 1-minute timings. At the end of the reading assessment, students were presented with a different copy of the reading passage and were required to identify the five cloze comprehension items deleted from the text.

**Peer Tutoring Training.** Target students were trained on peer tutoring earlier than the rest of the class. The purpose of this arrangement was to allow target students to receive more intensive instruction (i.e., working in pairs) than what they had been receiving in classroom instruction (i.e., small groups). Training for each target pair was conducted in the special education classroom. Because our target students had participated the previous year in a peer tutoring study (Al-Hassan, 2003), training for this study lasted only for one session. Students were able to perform all target tutoring skills correctly during training.

After all six target students had entered intervention, the rest of the class was engaged in peer tutoring. Prior to the class training, a meeting was held with the general education teacher to orient her to the peer tutoring program. The first author presented a scripted and timed training plan and explained all the steps of the activity. Training in the tutoring procedure involved demonstrations and role play. Moreover, the experimenter assisted the teacher with defining the huddle groups and their assigned huddle and tutoring areas.

In both trainings (i.e., target student pair training and total class training), the steps consisted of introducing to the students the purpose and meaning of peer tutoring (e.g., students can be good teachers, and everyone would learn how to teach each other words), modeling, and role-playing the skills. Training steps were based on the guidelines given by Cooke, Heron, and Heward (1983) in their peer tutoring handbook.

**Total Class Peer Tutoring.** At the beginning of every week and immediately before the intervention, the experimenter conducted a weekly pretest for each target student. The purpose of the pretest was to identify 10 sight words for each student. Five were the teacher's sight words, and five were unknown words pulled from other basic sight-word vocabulary lists (e.g., Dolch; Lowe, & Follman, 1974).

Peer tutoring sessions were conducted three times a week for 30 min each session. Each student was given a set of 10 sight words. Each target students' word set was determined during the weekly pretests. The nontarget students' (i.e., the rest of the students in the class) word set was decided by the classroom teacher. Peer tutoring sessions consisted of five components: tutor huddle, practice, testing, charting, and rewarding (Cooke et al., 1983).

The purpose of the *tutor huddle* component was to provide tutors with many opportunities to learn and practice the words that they would teach to their partners. Tutor huddle groups (with two to three tutors per group) were directed to designated spots in the classroom to practice the sight words that they would teach their tutees. A huddle period usually lasted for 4 minutes. During huddle time, students took turns reading their cards to the rest of the group. Each student held the card up and announced it, so that the rest of the group could see and hear the word. If the word was correct, huddle members would confirm the correct response by saying "yes."



If the word was incorrect, then the group would respond with "Say [the word]." In the case where no one knew a word or when there was a disagreement among the team, huddle members were advised to raise their hand and ask for either the teacher's or the experimenter's help. At the end of huddle time, students went to another assigned area to start the next phase of the intervention. The tutor huddle component was especially important for the present study, because each of the target students had a different set of sight words to learn.

The next component of peer tutoring was *practice*. Students joined their tutors to practice their words for 6 minutes. At the end of that time, students switched roles (i.e., tutors became tutees) and were given another 6 minutes to practice. The tutor presented the word cards one at a time and asked the tutee, "What word?" If the tutee responded correctly, then the tutor would provide praise and continue with the next card. If the tutee erred, then the tutor prompted his or her partner to try again. If the student responded incorrectly again or did not respond at all, the tutor would say the word (e.g., "Say *father*"). Tutors were encouraged to present word cards as many times as possible and to vary their social praise (e.g., "Great," "Super," "Fantastic," "Good job").

At the end of the practice period, students tested each other on the tutored words. The purpose of *testing* was to allow students to evaluate their learning on the mastered sight words. The tutor presented each word once. The tutor provided no verbal prompts, verbal feedback, or praise. If the student identified a word correctly, the tutor would place the word on the "happy face" pile. In the case of an incorrect response, the tutor would put the word card on the "X" pile. At the end of the testing, both students would record their word cards with either "happy face" (i.e., correct) or "X" (i.e., incorrect). At the end of the week, students placed all their cards in the "stop" pocket and a new set was given to them the next week. The purpose of having students practicing different words each week was to control for practice effects on the sight-word acquisition variable.

After testing, students counted the number of words identified correctly during testing and colored the same number of boxes on their chart. During each step of the session, both classroom teachers (general and special education) and the experimenter rewarded students (i.e., by placing stamps on their star cards) for demonstrating appropriate tutoring behaviors. On completion of the star card, students could exchange it with tangible items provided by the general education teacher. Furthermore, students who completed more than half of the boxes on their chart would get a "caught being good" coupon, on which they could write their name and enter it into a lottery. On completing the entire chart, students would again be eligible for other tangible rewards.

When the whole class participated in the peer tutoring activity, target students continued working with each other, whereas nontarget students were paired among themselves. The target students were tested individually at the end of each tutoring session. During initial assessment sessions, students

were tested on sight words three times a week, and reading fluency and comprehension were measured once a week using the DORF grade-level passages. In an effort to collect more data on the students' fluency, students completed the DORF passages two times a week starting from the seventh week of intervention. Nevertheless, as data collection continued, there was no clear evidence of the effects of peer tutoring on reading fluency based on the standardized DORF passages. Therefore, two additions were made during the assessment sessions:

1. Maintenance was incorporated as another dependent variable that would provide evidence of whether students' fluency was hampered by their inability to maintain the sight words.
2. Constructed paragraphs were incorporated to measure students' fluency and comprehension on tutored words.

**Posttest.** At the end of the treatment, the posttest included the same procedures as used in the pretest. That is, students were tested on the four subtests of the WJ-III.

### Social Validity

Three questionnaires were developed for assessing consumers' satisfaction with the intervention. Teachers (in both general and special education) and parents completed consumer satisfaction questionnaires at the end of the intervention. All classroom students were interviewed by a graduate student. The teacher questionnaire consisted of twelve 4-point Likert-type statements (i.e., from 1 = *strongly disagree*, to 4 = *strongly agree*) measuring the peer tutoring procedures (e.g., easiness, cost-effectiveness, impact of reward system) and the benefits of peer tutoring (e.g., improved sight-word acquisition, reading fluency, comprehension, promoting academic and social behaviors). There were also two open-ended questions where teachers were asked to provide additional thoughts or suggestions about the program.

The parent/guardian questionnaire consisted of seven statements based on a 4-point Likert scale and one open-ended question. Parents and guardians were asked to rate the importance of the intervention, appropriateness of procedures, and the advantages of peer tutoring on student's learning. In the one open-ended question, parents and guardians were given the opportunity to write further comments or thoughts about the program.

The student questionnaire included 12 questions. Using an ordinal scale of measurement from *liked very much* through *didn't feel anything* to *didn't like*, students expressed their opinion on six statements. Six open-ended questions were developed for measuring participants' thoughts and feelings about the components of the intervention.



## RESULTS

### Sight-Word Acquisition

All study participants learned more words during peer tutoring than during teacher-led classroom instruction. Students' sight-word acquisition performance is shown in Figure 1. Overall, target students as a group presented a mean of 3.97 (out of 5) correct words during baseline and 9.21 (out of 10) words during peer tutoring. Five of the six students demonstrated gains over classroom instruction. Erin, who was served under the learning disabilities category, demonstrated the highest mean percentage increase of 45.8 over baseline level. However, Irena was the only student with a mean percentage decrease of 0.9. Her high pretest ( $M = 4.4$  out of 5) and baseline ( $M = 4.87$  out of 5) scores did not permit her to exceed her previous performance during peer tutoring. During intervention, she was presented with more challenging words that were one to three grades above her grade level. Thus, her weekly pretest and intervention scores presented a mean of 2.89 and 9.65 (out of 10) correct words, respectively. Noteworthy are the group gains made during the last four weeks (Sessions 58–69) of intervention, when the classroom teacher did not present any sight words from her predetermined list. All target students started with a zero score in weekly pretests, and by the end of the week, they achieved a group mean of 9.36 (range = 8.9–9.8) correct words.

### Reading Fluency and Comprehension

All students increased their reading fluency and comprehension over classroom instruction levels on the DORF passages. The results are shown in Table 2. Target students demonstrated higher fluency gains on the constructed paragraphs than on the DORF passages during intervention. Fluency increases showed a group mean difference of 13.7 words per minute (wpm) between the constructed and the DORF passages. A smaller group mean increase of 3.9 wpm was found between baseline and intervention on the DORF passages.

Similar findings resulted for reading comprehension. Students achieved higher comprehension gains (mean increase = 1.1) on the DORF passages during peer tutoring. No substantial increase was evident between the constructed paragraphs and the DORF passages during intervention. Moreover, no comparisons can be made between baseline and experimental conditions for students' fluency and comprehension on constructed paragraphs, because these passages, as noted previously, had not been used from the beginning of the study.

### Maintenance

Maintenance assessments during and at the end of the study showed that all students but one retained high percentages of words practiced during peer tutoring. Specifically, group

mean retention of 87.4% (range = 72%–91.7%) was demonstrated during the study. Erin presented the lowest retention percentage (i.e., 72%).

A similar retention pattern was evident for the cumulative maintenance assessment. A group mean percentage of 87% (range = 66%–93.9%) was demonstrated. Again, Erin had the lowest retention rate of 66.2%. One reason for her low retention scores in both assessments was possibly her significant number of school absences. Although she showed the highest percentage gain in sight-word acquisition, her ability to maintain these words across time was hampered by her absences from peer tutoring sessions (16 out of 60).

### Pretest–Posttest Scores on Standardized Measures

The results of the standardized WJ-III subtests are shown in Table 3. Of the four subtests of the WJ-III, the students' performances increased most on the Letter–Word Identification and Word Attack subtests. Specifically, students evidenced 5-month and 7-month grade equivalent gains on the Letter–Word Identification and Word Attack subtests, respectively. Lower gains were obtained for Reading Fluency and Passage Comprehension, wherein students showed 1- and 3-month grade equivalent gains, respectively.

### Social Validity

The general and special education teachers “strongly agreed” that sight-word recognition is a critical skill for students to acquire for their future life. Both of them highly enjoyed participating in peer tutoring. They agreed that peer tutoring procedures were easy to implement, appropriate, and adequate, and that peer tutoring helped their students to actively engage in their learning. Teachers also agreed that peer tutoring improved their students' reading fluency and comprehension skills and that the reward system promoted students' academic and social performance. The general education teacher strongly agreed that peer tutoring improved her students' overall reading skills. She also responded that she would strongly recommend this program to other teachers, and she was planning to implement it in her class the following year. Furthermore, she noted, “This has worked very well for my special education students. The words are not always recognized in reading, but, overall, it is better that they learned the words. Special education students especially need repetition so that they can feel successful.” The special education teacher would also recommend this program to other teachers, and she was planning to implement it in her class the next year. She commented, “My students gained in their reading skills through peer tutoring. I enjoyed working with [the general education teacher].”

Parent/guardian questionnaires were given only to parents of the six target students. All six questionnaires were returned. Overall, parents “strongly agreed” that the peer tu-

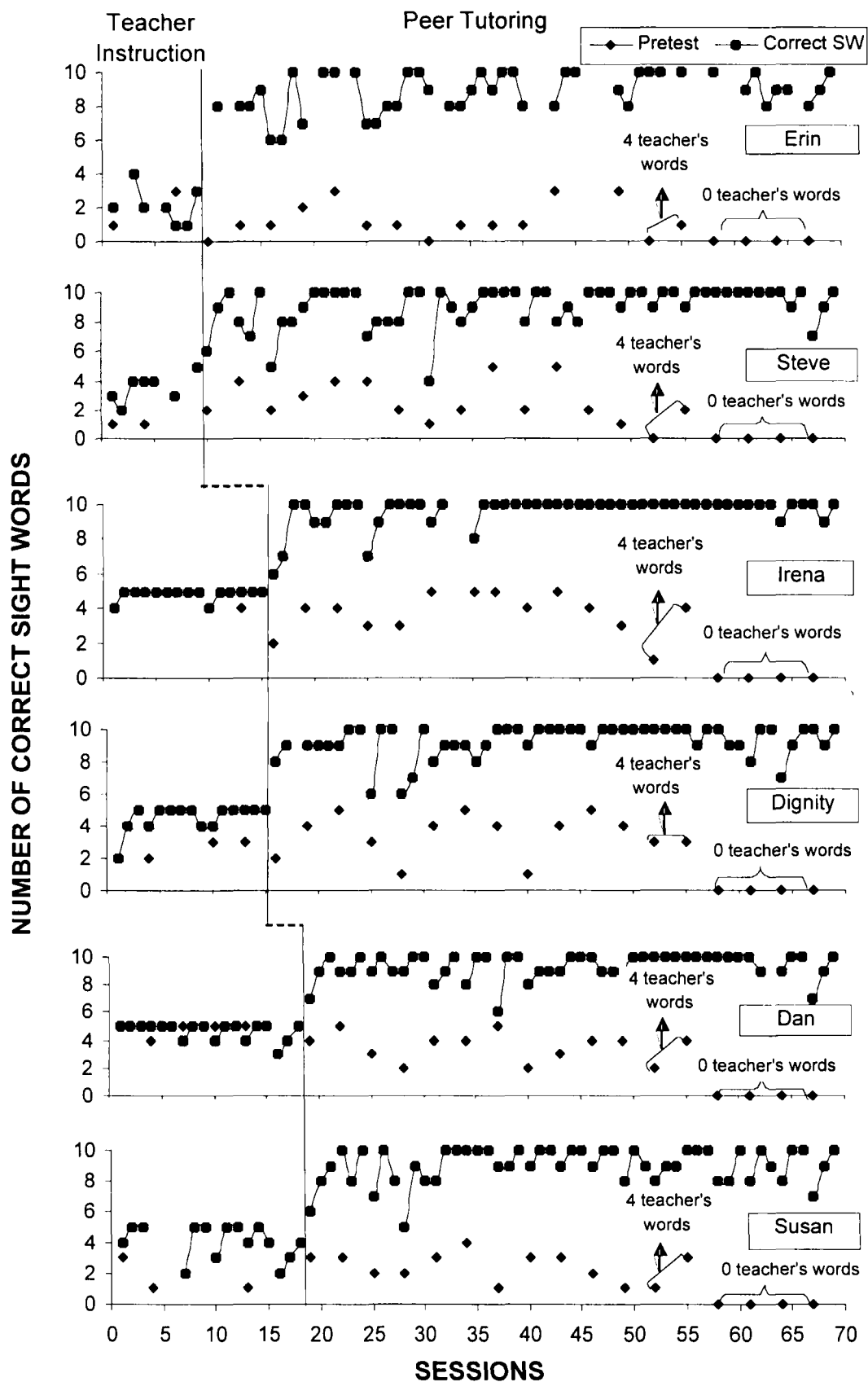


FIGURE 1. Sight-word acquisition during the intervention across the six target students. SW = sight words.

**TABLE 2. Mean Performance Scores on Reading Fluency (Words per Minute) and Comprehension for Target Students**

Student	Reading fluency			Reading comprehension		
	Classroom instruction		Peer tutoring	Classroom instruction		Peer tutoring
	DORF	DORF	CP	DORF	DORF	CP
Erin	14.7	16.5	28	2.3	3.9	4.4
Steve	20.5	23	36.4	3.0	3.8	4.6
Irena	24.6	35.4	54.1	2.8	4.2	4.4
Dignity	21.8	23.3	31.5	2.8	4.1	4.5
Dan	29.2	32.9	48.8	4.0	4.2	4.6
Susan	23.4	26.8	41.2	1.6	3.2	4.5
Group <i>M</i>	22.4	26.3	40	2.8	3.9	4.5

*Note.* DORF = *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) Oral Reading Fluency (Good & Kaminski, 2002) passages; CP = constructed paragraphs.

**TABLE 3. Pretest-Posttest Grade Equivalent Scores on the WJ-III Subtests for Target Students**

Student/pretest-posttest	Letter-Word Identification	Reading Fluency	Passage Comprehension	Word Attack
Erin				
10/08/03	1.6	< K.9	K.8	1.0
05/07/04	1.8	< K.7	1.4	1.6
Steve				
10/21/03	1.6	1.6	1.2	1.0
05/07/04	2.2	1.5	1.8	1.7
Irena				
10/10/03	1.9	1.3	1.6	1.9
05/07/04	3.0	2.0	2.0	2.4
Dignity				
10/24/03	1.9	1.8	1.4	1.0
05/07/04	2.4	1.6	1.7	2.3
Dan				
10/21/03	2.0	2.0	1.8	1.4
05/07/04	2.4	1.9	1.9	2.3
Susan				
10/08/03	1.9	1.0	1.4	1.0
05/07/04	2.3	1.2	1.5	1.0
Group <i>M</i>				
Pretest	1.8	1.4	1.4	1.2
Posttest	2.3	1.5	1.7	1.9

*Note.* WJ-III = *Woodcock-Johnson-III Tests of Achievement* (Woodcock, McGrew, & Mather, 2001).

toring helped their child to become a better reader. They noted that peer tutoring was important and beneficial for their child to learn more words and that their child explained the program in a positive manner. All parents said that they would like their child to continue participating in the peer tutoring program at school. Dignity's mother reported, "Dignity wasn't a good reader, but once getting started the peer tutoring, she now reads about the same level she should. I am very happy that they offer this at the school." Dan's mother also mentioned that "he needs more work but I can see the improvement."

Student interviews were conducted for all 14 students in the classroom. All students except one liked the peer tutoring very much. One student reported that she did not feel anything about peer tutoring because her partner was "mean." The majority of students (64.2%) liked very much being tutors, and a higher percentage (78.5%) of students preferred being tutees. All students liked their rewards very much. Participating students reported also that they liked saying different positive words to their partners (e.g., "good job," "terrific," "super"). Some of them (28.7%) would have preferred having easier but more words to practice during peer tutoring. No differences were found between the target and nontarget students' responses on the social validity measures.

## DISCUSSION

This study examined the effects of total class peer tutoring on sight-word acquisition, maintenance, reading fluency, and comprehension for six urban elementary-age students. The results of the study showed that five of the six participants increased their sight-word acquisition from teacher-led classroom instruction to total class peer tutoring. This finding is consistent with previous research (Al-Hassan, 2003; Giesecke & Cartledge, 1993; Heron et al., 1983) that clearly pointed out the effectiveness of this instructional strategy on low-achieving urban populations. Noteworthy is the fact that Erin and Susan, two of the five students that made gains over baseline, had been identified as students with disabilities (learning disabilities and attention-deficit/hyperactivity disorder). It is also worth noting that Erin, a target student with a disability, showed the highest mean percentage increase (i.e., 45.8%) during intervention. This lends support to the position that total class peer tutoring can be a viable instructional tool for the inclusion of students with disabilities in the least restrictive settings. As Heron, Welsch, and Goddard (2003) noted, "tutoring can provide a 'boost' to students with disabilities by providing an opportunity for them to contribute to the overall general education instruction in the classroom" (p. 297).

Although this study targeted only 6 of the 14 students in the joint classroom, data were collected for the 8 other students by the general education teacher. Based on anecdotal reports and informal discussions between the experimenter

and the teacher, all 8 students increased their sight-word acquisition during total class peer tutoring. It should be noted that the general education teacher introduced only five words in her daily instruction. After intervention had started, the teacher increased the number of words to 10. At the end of the study, both general and special education teachers were satisfied with the performance of their students, especially of the students with disabilities, as they made up the majority (57%) of students in the class. Therefore, there is sufficient evidence that under well-structured, closely organized conditions, low-achieving students with and without disabilities can learn effectively.

Nevertheless, there was one target student who presented a minor mean percentage decrease (0.9%) on sight-word acquisition by the end of the study. Irena's case is a reflection of the teacher's low expectations and the teacher's lack of differentiated instruction toward urban students. Specifically, when looking at Irena's graph, one can clearly observe that her weekly baseline pretest scores were very high ( $M = 88\%$ ). This means that Irena already knew the five teacher words before the teacher taught them to the class. Hence, her performance across baseline sessions was quite high (97.4%). When the intervention started, Irena's weekly pretests ( $M = 30\%$ ) showed that she could read her second-grade sight words. Therefore, she was presented with more challenging words, one to three grade levels above her grade placement. Despite the fact that she acquired these words ( $M = 96.5\%$ ) during peer tutoring sessions, the more difficult and higher grade level sight words did not permit her to exceed her very high baseline levels, which were established with much easier words. Teachers' low expectation is an inhibitory academic achievement factor for urban students (Cartledge, 2002). Lowered expectations and instructional goals, common in inner-city schools, need to be points of critical focus in interventions for urban students.

Data on maintenance probes indicated that the majority (83.3%) of the target students were able to maintain their learned words after 2 to 3 weeks of intervention as well as at the end of the study (i.e., after 4 months). Students' high retention percentages ( $M = 87.5\%$ ; range = 72%–91.7%) lend support to the findings of Heron et al. (1983), as students seemed to benefit from peer tutoring procedures.

Of interest are the findings on students' reading fluency and comprehension. The data on standardized DORF passages showed that students' fluency and comprehension increased from baseline to intervention. However, moderate effects were evident, and no functional relationship could be determined between peer tutoring and students' fluency or comprehension based on the DORF passages. Students did not increase their performance (i.e., fluency and comprehension) substantially, and one reason for this might have been the fact that these passages were an indirect generalization measure of students' tutored words. In other words, the DORF passages contained very few of the sight words that the students had practiced during peer tutoring. Thus, the stu-

dents' reading performance on the DORF passages did not improve significantly.

On the other hand, higher reading rates were evident from the constructed paragraphs, which were specifically developed to include the words that students were learning in the peer tutoring. Students increased their wpm by a group average of 14 wpm compared to the DORF passages during intervention. Students' higher fluency on constructed paragraphs could be explained as the result of a more direct generalization measure of students' tutored words. Constructed paragraphs included at least 20 words that students had practiced and learned during peer tutoring. Therefore, students were able to generalize (i.e., identify and read) these words in context. This finding lends support to other studies on peer tutoring (Al-Hassan, 2003; Barbetta et al., 1991) in which tutored words could be generalized when presented in context (i.e., sentences). This is one of the few studies, however, that also measured reading fluency. Although we recognized that there are more direct ways to develop reading fluency (e.g., repeated readings), these findings show that sight-word acquisition can also contribute to more fluent reading with understanding. That is, fluency in reading isolated sight words can contribute to fluency in connected text. Students' reading comprehension on constructed paragraphs was higher than on the DORF passages. However, the difference was minimal, because students were able to identify at least four out of five cloze comprehension items during intervention on either the DORF passages or the constructed paragraphs.

Students' pretest and posttest scores on the standardized WJ-III subtests showed that students made the highest gains in Letter-Word Identification and Word Attack subtests. As the intervention focused on learning sight words, the students' gains on these two subtests of the WJ-III were no surprise. Once again, this finding underscores the beneficial effects of the peer tutoring.

In light of the aforementioned findings, several study limitations arise. First, no reading fluency and comprehension measures were taken on constructed paragraphs during baseline. Hence, no comparisons or functional relationship can be discussed between the teacher-led instruction condition and peer tutoring relative to the constructed paragraphs measure. Therefore, future studies might address this direct generalization measure at the beginning of the study. Second, students stopped practicing their set of sight words by the end of the week, and a new set was given at the beginning of next week. This limited their opportunity to benefit fully from the peer tutoring practice to master all 10 sight words before moving on to different words. As suggested in the research literature (e.g., Cooke et al., 1983), a criterion of three consecutive successes is recommended before removing a word, thereby increasing the likelihood of maintenance. Third, students' absences—a common problem in urban settings—presented special difficulties in accomplishing the intervention goals. One of the target students was absent for 30% of the tutoring sessions. Her absences might have affected her low retention rate and her lower integrity scores on tutoring

behaviors. Fourth, several word lists were used in the study. However, several common words were found among the lists, which created some difficulties during weekly pretests. Despite the experimenter's efforts to avoid presenting the same words during weekly pretests, the possibility of introducing the same word more than once did occur. These duplications were accounted for in the assessments, so that students' results reflect the actual number of words learned and retained. For future purposes, it would be easier to develop one master word list, from which words can then be taken for pretesting. Finally, no comparisons can be made between baseline and intervention with respect to the maintenance variable, because maintenance was only assessed during intervention. The inclusion of maintenance for longer time periods from the beginning of the study would have remedied this problem.

Despite the number of limitations and logistical problems, the results of this study suggest several reasons why teachers should implement total class peer tutoring. First, total class peer tutoring is a field-tested instructional activity that provides increased student opportunity for successful correct responses and high retention rates. Second, given the ease of training students as tutors, teachers can take advantage of this rich available resource pool in their classrooms. Thus, teachers have the opportunity to maximize their instructional influence on the classroom as well as to provide individualized instruction. Third, peer tutoring emphasizes important social skills that teachers often overlook. Such social skills include behaviors such as making positive statements to others, giving and accepting feedback, taking turns, and cooperating with peers. To the extent that peer tutoring is established in the classroom, teachers can encourage the generalization and maintenance of these skills across other instructional activities. Finally, given the positive consumer satisfaction outcomes and the simple and easy tutoring procedures, total class peer tutoring should be integrated into teachers' preservice training. Guiding preservice teachers in structuring peer tutoring, providing appropriate training, and monitoring and evaluating their progress will ultimately increase the likelihood of their continued use of such strategies throughout their teaching career.

In summary, this study showed that total class peer tutoring is a viable instructional alternative for both general and special education teachers to support and meet the reading needs of minority and culturally diverse students. Students increased their sight-word knowledge and generalized these acquisitions to contextual passages where they showed increases in fluency and comprehension. Future directions in research might continue addressing the limitations noted in the present study to provide classroom teachers with an even more efficient instructional tool. ■

---

LEFKI KOUREA, MA, is a doctoral student in the Special Education Program at The Ohio State University. Her current interests include effective classroom management strategies and academic and behavioral interventions for students with behavioral disorders. GWENDOLYN CARTLEDGE, PhD, is a professor in special education at The Ohio State University. Her research and teaching have focused on students with learning and behavior

disorders, social skills, preventive interventions, and urban education. **SHOBANA MUSTI-RAO**, PhD, is an assistant professor in special education at the University of Cincinnati. Her areas of interest include preventive interventions (academic and behavioral) for urban learners identified as being at risk for school failure. She is particularly interested in strategies that follow the principles of universal design for learning. Address: Gwendolyn Cartledge, School of Physical Activity and Educational Services, The Ohio State University, 356 Arps Hall, 1945 N. High Street, Columbus, OH 43210; e-mail: cartledge.1@osu.edu.

## AUTHORS' NOTE

Preparation of this manuscript was supported by the U.S. Department of Education Model Demonstration Project for Children with Disabilities Grant no. H324T010057.

## REFERENCES

- Al-Hassan, S. (2003). *Reciprocal peer tutoring effect on high frequency sight word learning, retention, and generalization of first- and second-grade urban elementary school students*. Unpublished doctoral dissertation, Ohio State University, Columbus.
- Arreaga-Mayer, C. (1998). Increasing active student responding and improving academic performance through classwide peer tutoring. *Intervention in School and Clinic*, 34, 89-95.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91-97.
- Barbetta, P. M., Miller, A. D., Peters, M. T., Heron, T. E., & Cochran, L. L. (1991). TUGMATE: A cross-age tutoring program to teach sight vocabulary. *Education and Treatment of Children*, 14, 19-37.
- Brady, N. C. (1997). The teaching game: A reciprocal peer tutoring program for preschool children. *Education and Treatment of Children*, 20, 123-149.
- Butler, F. M. (1999). Reading Partners: Students can help each other learn to read! *Education and Treatment of Children*, 22, 415-426.
- Cartledge, G. (2002). Issues of disproportionality for African American learners. In L. Bullock & R. A. Gables (Eds.), *Culturally and linguistically diverse students with behavioral disorders* (pp. 15-24). Arlington, VA: Council for Children with Behavioral Disorders.
- Cooke, N. L., Heron, T. E., & Heward, W. L. (1983). *Peer tutoring: Implementing classwide programs in the primary grades*. Columbus, OH: Special Press.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.
- Delquadri, J. C., Greenwood, C. R., Stretton, K., & Hall, V. R. (1983). The Peer Tutoring Spelling Game: A classroom procedure for increasing opportunity to respond and spelling performance. *Education and Treatment of Children*, 6, 225-239.
- DuPaul, G. J., & Henningson, P. N. (1993). Peer tutoring effects on the classroom performance of children with attention deficit hyperactivity disorder. *School Psychology Review*, 22, 134-142.
- Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of Educational Psychology*, 81, 173-177.
- Ferri, B. A., & Connor, D. J. (2005). In the shadow of Brown: Special education and overrepresentation of students of color. *Remedial and Special Education*, 26, 93-100.
- Giesecke, D., & Cartledge, G. (1993). Low-achieving students as successful cross-age tutors. *Preventing School Failure*, 37(3), 34-43.
- 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. Retrieved October 10, 2003, from <http://dibels.uoregon.edu>
- Gottlieb, J., Alter, M., Gottlieb, B. W., & Wishner, J. (1994). Special education in urban America: It's not justifiable for many. *The Journal of Special Education*, 27, 453-465.
- Grant, P. L. (1979). The cloze procedure as an instructional device. *Journal of Reading*, 22, 699-705.
- Greenwood, C. R., Arreaga-Mayer, C., Utley, C. A., Gavin, K. M., & Terry, B. J. (2001). Classwide peer tutoring learning management system: Applications with elementary-level English language learners. *Remedial and Special Education*, 22, 34-47.
- Greenwood, C. R., & Delquadri, J. (1995). Classwide peer tutoring and the prevention of school failure. *Preventing School Failure*, 39(4), 21-25.
- Heron, T. E., Heward, W. L., Cooke, N. L., & Hill, D. S. (1983). Evaluation of a classwide peer tutoring system: First graders teach each other sight words. *Education and Treatment of Children*, 6, 137-152.
- Heron, T. E., Welsch, R. G., & Goddard, Y. L. (2003). Applications of tutoring systems in specialized subject areas: An analysis of skills, methodologies, and results. *Remedial and Special Education*, 24, 288-300.
- Heward, W. L., Heron, T. E., & Cooke, N. L. (1982). Tutor huddle: Key element in a classwide peer tutoring system. *The Elementary School Journal*, 83, 114-123.
- Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis*, 10, 141-150.
- Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis*, 27, 49-61.
- Lazerson, D. B. (1980). "I must be good if I can teach!"—Peer tutoring with aggressive and withdrawn children. *Journal of Learning Disabilities*, 13(3), 43-48.
- Lo, Y., & Cartledge, G. (2004). Total class peer tutoring and interdependent group oriented contingency: Improving the academic and task related behaviors of fourth-grade urban students. *Education and Treatment of Exceptional Children*, 27, 235-262.
- Lowe, A. J., & Follman, J. (1974). Comparison of the Dolch list with other word lists. *Reading Teacher*, 28(1), 40-44.
- Maheady, L., & Harper, G. F. (1987). A class-wide peer tutoring program to improve the spelling test performance of low-income, third- and fourth-grade students. *Education and Treatment of Children*, 10, 120-133.
- Maheady, L., Sacca, M. K., & Harper, G. F. (1987). Classwide student tutoring teams: The effects of peer-mediated instruction on the academic performance of secondary mainstreamed students. *The Journal of Special Education*, 21, 107-121.
- McDonnell, J., Mathot-Buckner, C., Thorson, N., & Fister, S. (2001). Supporting the inclusion of students with moderate and severe disabilities in junior high school general education classes: The effects of classwide peer tutoring, multi-element curriculum, and accommodations. *Education and Treatment of Exceptional Children*, 24, 141-160.
- Miller, L. J., & Kohler, F. W. (1993). Winning with peer tutoring. *Preventing School Failure*, 37(3), 14-19.
- Miller, A. D., Barbetta, P. M., & Heron, T. E. (1994). START tutoring: Designing, training, implementing, adapting, and evaluating tutoring programs for school and home settings. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, et al. (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 265-282). Monterey, CA: Brooks/Cole.
- Nazzari, A. (2002). Peer tutoring and at-risk students: An exploratory study. *Action in Teacher Education*, 24(1), 68-80.
- Simmons, D. C., Fuchs, D., Fuchs, L. S., Hodge, J. P., & Mathes, P. G. (1994). Importance of instructional complexity and role reciprocity to classwide peer tutoring. *Learning Disabilities Research & Practice*, 9, 203-212.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Examiner's manual: Woodcock-Johnson-III Tests of Achievement*. Itasca, IL: Riverside.
- Wright, J., Cavanaugh, R. A., Sainato, D. M., & Heward, W. L. (1995). Somos todos ayudantes y estudiantes: A demonstration of a classwide peer tutoring program in a modified Spanish class for secondary students identified as learning disabled or academically at-risk. *Education and Treatment of Children*, 18, 33-52.