Using Peer-Mediated Constant Time Delay to Teach Content Area Vocabulary to Middle School Students with Disruptive Behavior

Rachel Wannarka
Kathy Ruhl
Richard Kubina

ABSTRACT: Vocabulary knowledge is essential to reading comprehension and general academic success for all learners. For students with disruptive behaviors, vocabulary demands encountered in secondary settings and content area texts can be especially challenging in part due to the frequent occurrence of unfamiliar vocabulary. It is critical that vocabulary instruction make judicious use of time as students move rapidly through curriculum. This article presents results of a study investigating a peer-mediated constant time delay intervention to teach science and social studies vocabulary to middle school students with teacher-identified disruptive behavior. Students demonstrated efficient mastery of target definitions and learned definitions incidentally as a result of teaching their peers. Students also demonstrated generalization to word reading, word production, and to a lesser degree recognition of correct usage on a multiple-choice measure. Implications for practice include selecting useful target words, aligning instructional methods and goals, utilizing carefully selected peer arrangements, and providing sufficient training.

THE PROBLEM

Students with disruptive behavior often experience a host of negative academic and social outcomes that hinder their school success. Whether or not they have identified disabilities such as learning disabilities (LD) or emotional/behavior disorders (EBD), students who display substantial disruptive behaviors tend to lag behind their

peers academically (Heward, 2009). Reading problems are a special concern because such a large degree of academic learning is language dependent and relies on reading comprehension (Baker, Simmons, & Kameenui, 1995). It is well documented that vocabulary knowledge is a critical component of reading comprehension (Baumann & Kameenui, 1991; NRP, 2000) and a prerequisite for fluent reading (Joshi, 2005). The correlation between vocabulary knowledge and reading comprehension has been demonstrated across all grade levels and in many countries (Joshi, 2005).

As students progress through grade levels, the amount and complexity of information they are exposed to increases dramatically (Bulgren, Deshler, Schumaker, & Lenz, 2000) and curricula increasingly rely on vocabulary (Harmon, Hedrick, & Wood, 2005). Secondary students face particular difficulties in their content area classes because such a large proportion of critical information comes from textbooks that tend to be poorly organized and contain a substantial amount of unfamiliar, technical vocabulary (Mastropieri, Scruggs, & Graetz, 2003). Because students with disruptive behaviors are likely to be struggling readers, their difficulties with content area textbooks may tend to be more severe. Teachers have a variety of tools to make these texts more understandable such as graphic organizers and study guides, but vocabulary knowledge is necessarily a limiting factor to the effectiveness of such external supports.

In inclusive settings it is important to develop interventions that not only assist students with disruptive behaviors (whether or not they have been identified as having EBD), but also their peers without these issues. Teachers are responsible for the education of all their students and are more likely to use a particular instructional method if it is appropriate for the range of students in their classes.

Instructional efficiency is critical for students who are behind their peers academically. Regardless of the reason for lower academic performance, in order to be successful in school, students who are behind their peers need effective instruction presented in an efficient manner. In the area of vocabulary, judicious use of instructional time is paramount due to characteristics of typical textbooks with high quantities of unfamiliar vocabulary needed to understand the content. For students with disruptive behavior or special needs who may have substantial vocabulary and general reading deficiencies, the need for expeditious vocabulary instruction is clear.

One consideration in efficiency is the level of mastery to which terms should be taught. Words students will need to use in their expressive

vocabularies (i.e., writing and speaking) require more fluency than words students need only to understand within a limited arena. Indeed many literacy researchers (e.g., Baker et al., 1995; Jitendra, Edwards, Sacks, & Jacobson, 2004; NRP, 2000) agree that a fundamental underpinning of well-designed vocabulary instruction is alignment of academic goals and instructional methods. Words that occur frequently across multiple contexts or have several common meanings should be taught so that students are able to use the words appropriately in their expressive vocabularies (Kameenui, Simmons, & Darch, 1987; Stahl & Fairbanks, 1986) level. Standards for utility in the receptive vocabulary alone, however, are substantially less stringent (Baumann & Kameenui, 1991) and may be met by association-level (i.e., between a word and a single definition or synonym) knowledge. Reading requires use of receptive vocabulary (students do not have to generate terms, only understand them), thus it may not be necessary and may, in fact, be wasteful to invest large amounts of teacher and student effort with in-depth instruction. For example, in order to understand a selection from a secondary science text, students need only to have an association-level understanding of words that are highly technical (e.g., oxidize, parallax) and unlikely to occur in other contexts.

Reducing the vocabulary gap between good and poor readers is a broad and complicated task that is rightly the responsibility of all teachers. However, the large amount of unfamiliar and technical vocabulary in textbooks is a specific barrier to success in science, social studies, and other content area classes. Overcoming this barrier by efficiently increasing word learning may prove simpler than broad-based vocabulary gap reduction because these words do not need to be learned deeply for comprehension to improve (Baumann & Kameenui, 1991).

Researchers have examined a wide range of vocabulary instruction methods, such as mnemonic strategies (Mastropieri, Scruggs, & Fulk, 1990), cognitive strategies such as semantic mapping (Bos & Anders, 1990), activity-based methods (Scruggs, Mastropieri, Bakken, & Brigham, 1993), and various forms of computer-assisted instruction (Koury, 1996). Generally speaking, any form of vocabulary instruction improves word learning compared with no instruction or the "dictionary method" in which students write definitions from word lists. Constant time delay (CTD) was selected as the method of instruction for this study for two reasons: most other methods rely on more time-intensive procedures (e.g., mnemonic strategies) or require specialized equipment (i.e., computer-assisted instruction), and CTD is designed to minimize student errors by providing immediate correction.

PREVIOUS RESEARCH

In terms of vocabulary instruction, CTD is a response prompting strategy with two defining characteristics: (1) initial trials involve presentation of the vocabulary word on a flash card followed immediately by the definition, and (2) subsequent trials involve presenting the vocabulary word followed by a fixed amount of time (e.g., 3 to 5 seconds) during which the student may respond by supplying the definition. If the student correctly states the definition, reinforcement is provided; the definition is stated as error correction if the student fails to do so (Wolery et al., 1992).

Because the fundamental reason for academic reading is comprehension and association-level vocabulary knowledge is largely sufficient for maximal comprehension (Baumann & Kameenui, 1991), CTD may be effective and efficient for students with disruptive behavior as an intervention to improve content area text comprehension. Research has shown CTD procedures to be effective and efficient compared with other instructional methods used to teach academic skills (McDonnell, Johnson, Polychronis, & Risen, 2002; Schuster, Stevens, & Doak, 1990) and to have low error rates (Wolery, Ault, & Doyle, 1992).

Existing CTD vocabulary studies (e.g., McDonnell et al., 2002; Schuster et al., 1990) indicate that error rates and efficiency are superior when compared to other methods, but both the number of investigations and range of students represented are small. Schuster and colleagues (1990) utilized a CTD procedure to teach content area vocabulary words to upper-elementary students with LD, finding that the three participants demonstrated mastery of target definitions (mean length of six words) after eight or nine brief instructional sessions. CTD has also been used successfully in a class-wide peer tutoring format to teach vocabulary definitions to sixth graders with and without LD with high efficiency and low error rates (Hughes & Fredrick, 2006).

Peer-mediated instruction in many forms (e.g., peer-assisted learning strategies [Fuchs, Fuchs, Kazdan, & Allen, 1999], class-wide peer tutoring [Greenwood, Arreaga-Mayer, Utley, Gavin, & Terry, 2001], and reciprocal teaching [Palincsar & Herrenkohl, 2002]) has a strong research base of support in content area classes (Kroeger, Burton, & Preston, 2009). To ensure that limited class time is used well, literacy experimenters have contended that future investigations should address the academic and social or emotional benefits of peer tutoring interventions for students with EBD (e.g., Spencer, Scruggs, & Mastropieri, 2003). In an analysis of peer-mediated interventions on academic

outcomes for students with EBD, Ryan, Reid, and Epstein (2004) found large positive effect sizes for adolescents across content areas.

THE SOLUTION

As previously noted, a primary rationale for using CTD interventions is that they tend to result in more rapid learning than other methods. An additional way to optimize available instructional time with CTD may be to capitalize on opportunities for incidental learning by having students work on learning definitions in pairs, with each learner focusing on a different set of definitions. Promoting opportunities for incidental learning permits more information acquisition in the same amount of instructional time. Paired instructional arrangements may be especially useful for students with social difficulties that commonly accompany behavior problems and special needs because these arrangements allow structured opportunities for students to engage in appropriate peer interactions.

This article describes the results of a study of a peer-mediated CTD procedure to teach content area vocabulary definitions to middle school students identified by their teachers as being "exceptionally disruptive." The peer format allowed incidental learning to be examined as well. Generalization measures assessed were word reading, word production upon definition presentation, and recognition of correct usage on a written multiple-choice test.

EVIDENCE OF EFFECTIVENESS

A multiple probe design (Tawney & Gast, 1984) was used in this study. This variation of the multiple baseline design maintains the experimental control of multiple baseline studies by sequentially introducing the independent variable (in this study, CTD vocabulary instruction) to one participant at a time while baseline data were collected intermittently rather than continuously on behaviors yet to be introduced. This has the benefit of minimizing negative behaviors such as noncompliance resulting from frustration that may occur when students are repeatedly asked to perform a skill they have yet to learn, especially for students who receive intervention toward the end of a study because they would spend the most time in the baseline

phase of the intervention. An additional benefit of multiple probe designs is that practice effects are minimized.

Setting

The study was implemented in an urban charter school of 320 students in Grades 5 through 8 in a large Northeastern state. Forty-one students had identified special needs and the principal reported that approximately 15 students without identified disabilities had behavior problems substantial enough to impact their learning and that of their classmates. Eighty percent of the student population was African American and 20% was Hispanic. Ninety percent of the student body was receiving federal free or reduced-price lunches and gender distribution was roughly equal. Instructional and probe sessions took place in vacant rooms around the school (e.g., library, conference rooms).

Participants

Participants were six sixth-grade students identified by at least two teachers as displaying substantially disruptive classroom behavior that was outside the norms of typical and expected school behavior (e.g., shouting across the room during class, fighting). All sixth-grade teachers were asked to nominate five exceptionally disruptive students. Students identified by at least two teachers were put on a list which was given to the sixth-grade science and social studies teachers. These teachers were asked to identify students from that list who were performing academically in the bottom quartile in their classes, and nine students were selected by both teachers. Of those nine, six were randomly selected for inclusion in the study and all six agreed to participate and returned signed consent forms.

All participating students were assessed to ensure they possessed sufficient verbal imitation skills to participate in the study. For the verbal imitation assessment, 15 phrases of length and structure comparable to the target definitions were selected and the researcher read each phrase to each student individually and instructed him or her to repeat it. All students did this with 100% accuracy.

Four girls and two boys participated in the study. All were African American and all qualified for free lunch. Tamika was 13 years old with a fourth-grade reading level at the beginning of the study and had no identified special education needs. Keisha was 12 years old

and diagnosed with an LD in the area of reading comprehension and a fourth-grade reading level. Jamila was 11 years old with a sixth-grade reading level at the time of the study and had no identified special educational needs. Jatara was 12 years old with a fourth-grade reading level and no identified special educational needs. Jamar was 12 years old with an identified LD in the area of reading comprehension and had a reading level of 4.6 (i.e., between fourth and fifth grade). Keon was 12 years old, had no identified special educational needs, and had a seventh-grade reading level. Reading levels were assessed using the Developmental Reading Assessment. There was no indication in any of the students' files that they had been assessed for possible language deficits (e.g., specific language impairment), although Keisha was receiving weekly speech therapy to address articulation problems. With the exception of Jamar who was new to the school and had his reading level determined by his previous school, all students' reading levels were assessed by the sixth-grade reading teacher at the start of the school year.

Three pairs were formed and each pair had two stable baseline sessions at the beginning of the study. The first pair to engage in the intervention phase had a third baseline session prior to intervention, and each subsequent pair had two baseline sessions before beginning their intervention phase in order to establish that rates of correct responding were consistently at 0%. Both members of each pair had to reach criterion before the next pair began the intervention phase.

Materials

Word determination flashcards. Potential science and social studies target words were handwritten on one side of index cards and short definitions devised by the researcher and approved by the students' science or social studies teacher were written on the back. Twenty words in each content area were used to create a deck of 40 cards.

Training flashcards. A stack of 10 flashcards to be used to teach participants to conduct the CTD intervention was created by taking sample words from a geometry text. One target vocabulary word was written on the front of each card, and a simplified glossary definition with three underlined key words was written on the back.

Intervention flashcards. A stack of flashcards containing each target word three times (21 total cards) was created for each content area. Target words were written on the front and definitions were written on the back. Three key words or short phrases per definition were underlined

in order to provide a simple way for students to approximate whether or not their partners were able to produce a correct definition.

Correct usage recognition generalization assessment. A 14-item multiple-choice test was developed in which each target word was the subject of one item. Each item consisted of the target word above four sentences: one in which the target word was used correctly and three in which it was used incorrectly. Sentences were labeled A through D and the position of the correct choice was randomized across items.

Dependent Variables

Target word definitions. The primary dependent variable was the percentage of correct target definitions stated during instructional and probe sessions in the main phase of the study and for probe sessions during maintenance checks. The percentage correct was calculated by dividing the number of correct responses by 21, the number of trials in each session. During instructional sessions, tutors marked their partners' responses as correct if the tutee started to speak within 5 seconds of word presentation and provided a definition that included all underlined words.

Other variables of interest measured generalization and were all assessed prior to beginning the intervention and again once pairs reached mastery on target word definitions. All generalization measures were collected by school staff with the exception of pre-test measures for Tamika and Keisha which were collected by the researcher and observed by school staff.

Nontarget word definitions. The percentage of correct definitions stated after visual presentation of nontarget words (i.e., each student's partner's target words upon which the student received no direct instruction) was also recorded to serve as a measure of incidental learning.

Word production. This was a measure of the percentage of correct target words stated by participants after listening to definitions. Stating a word is a less cognitively demanding (although still useful) task than stating a definition, and word stating has been used as a primary dependent variable in at least one other CTD vocabulary intervention (Hughes & Fredrick, 2006). Procedures were identical to target word definition procedures except that school staff showed and read the definition side of the flash card to students as the prompt.

Word reading. This was a measure of the percentage of target and nontarget words read out loud correctly upon visual presentation of the words. To assess word reading, staff presented each word on a

flashcard to each student and recorded whether the student was able to pronounce the word correctly.

Correct usage recognition. This was a measure assessing the number of correct responses on a researcher-developed multiple-choice exam in which students had to identify correct usage of target and non-target vocabulary. Number rather than percentage was the measure of interest on the multiple-choice exam because the total number of opportunities to respond was fewer than 20 (Tawney & Gast, 1984), and small changes in the number correct could translate to relatively large and potentially misleading changes in percentage.

For the multiple-choice generalization exam, effort was made to create a set of incorrect choices that each varied one (and only one) of the underlined words or phrases for all terms. For example, the definition for the social studies target word *refugee* read, "A person who leaves their country during a war for safety." One incorrect choice depicted a person moving across town (varying "country"), another described a person moving during the hurricane season (varying "war"), and the third represented a person moving to save money (varying "safety").

Due to the nature of definitions for the words *xenophobia, monopoly, assimilate, cell, trachea,* and *pathogen,* it was not possible to construct incorrect choices in this manner for each target word. For example, the definition for the science target word *trachea* read "The tube in the throat that moves air to and from the lungs," and a choice depicting something other than air moving to and from the lungs would be incorrect, but it would also be nonsensical and would thus increase the chances of a correct response without the corresponding knowledge the item was meant to assess. Directions read, "Circle the letter of the sentence that shows an example of each underlined word." An example item appears below.

Refugee

Marnie and her mother went to live with cousins in another country during the war because they wanted to save money.

Marnie and her mother went to live with cousins on the other side of town during the war because they didn't want to get hurt.

Marnie and her mother went to live with cousins in another country during the war because they didn't want to get hurt.

Marnie and her mother went to live with cousins in another country during the hurricane season because they didn't want to get hurt.

In this example item, choice (a) varies "safety," (b) varies "country," (d) varies "war," and all are incorrect. Choice (c), the correct option, is the only one that does not vary any key definitional elements.

Procedures

Independent variable. The independent variable was a peer-delivered CTD vocabulary intervention. Participants were assigned partners and provided instruction to their partners on seven definitions from either science or social studies content areas, and they received instruction from their partners on seven definitions from the other content area.

Target word selection. Twenty potential target words in each content area were generated by the students' science and social studies teachers based on importance to instructional goals, content specificity of definitions, and perceived likelihood that the definitions would be unknown to all students. The researcher devised short, functional definitions of all potential words and the teachers approved the definitions as capturing the essential ideas as they would be used in future lessons. Forty flashcards with a potential word on one side and the corresponding definition on the reverse were created for target word selection.

In order to verify words that were unknown, each flashcard was displayed to participants individually by the researcher who asked students to verbally provide a definition. If a student was unable to produce an approximate definition within 15 seconds for two consecutive trials, the word was included as a candidate for training. Words deemed unknown by all participants were separated and seven words per content area were randomly selected for instruction.

Intervention training. The experimenter trained one special educator and one administrative staff member in the procedures so they would be able to monitor the CTD intervention. First, the experimenter provided a brief overview of the theoretical underpinnings of CTD in academic settings and described the study procedures including error correction and data sheet usage. Next, the experimenter acted as the tutor and used the practice set of geometry vocabulary flashcards to conduct a session with each school staff member acting as the tutee. Then the school staff participated in two instructional sessions, playing the role of tutor and tutee once each. The experimenter provided corrective feedback as necessary about marking data sheets, waiting the appropriate amount of time between word presentation and initiation of correction procedures, and provision of verbal confirmation in response to correct definitions. Neither staff required any correction

after the first 5 of the 21 trial sessions. Staff were given the opportunity to ask questions during and after the training.

Students were assigned to pairs based on schedules and with their homeroom teachers' input in order to minimize potential behavior problems. One student per pair was randomly assigned to science target words and one to social studies target words. Students were paired as follows: Tamika (science target words) and Keisha (social studies); Jamila (science) and Jatara (social studies); and Keon (science) and Jamar (social studies).

Participants were trained with the practice word set in their pairs 1 to 2 days before they were to begin instruction. The experimenter trained the first pair while the two staff members observed. School staff members conducted intervention training for the remaining pairs. Participants were instructed to show target words to their partners and read them out loud when acting as tutors. The adult (i.e., school staff or researcher) supervising the instructional sessions provided corrective feedback on word pronunciation if necessary.

Instructional sessions. Each instructional session consisted of 21 trials in which the seven target words were presented three times. Flashcard decks were shuffled prior to each session to randomize word order. Because target words were unknown to tutors as well as tutees, tutors could not be expected to independently judge whether tutees produced a functional definition, so tutors marked their partners' responses as correct during instructional sessions if the tutee provided a definition that included all underlined words. The tutee also had to start speaking within 5 seconds of word presentation for a response to be marked correct.

For each pair's first instructional session, tutors were directed to read the definition on the back of the flashcard and wait for their partners to repeat it. For all subsequent sessions, tutors were instructed to silently count to five ("one-one thousand, two-one thousand," etc.) while waiting for their partners to begin saying the definition. If tutees produced a definition containing all three underlined words or phrases, tutors were instructed to inform their partners that the definition was correct and to mark it as such on the data sheet. If tutees did not begin to speak within five seconds or failed to reproduce all three underlined words or phrases, tutors were instructed to mark the response as incorrect, read the definition from the back of the flashcard, and wait for their partners to repeat the definition. Corrections were repeated if tutees failed to repeat target definitions accurately. Criterion was 100% correct responses for three consecutive

trials. If one participant reached criterion before his or her partner, instructional sessions consisted of only the student who had yet to reach criterion-receiving instruction.

There were no occasions on which students had more than one instructional session per day. Number of sessions per week ranged from three to five depending on school schedules and student attendance.

Probe sessions. Probe sessions were conducted by trained school staff or the experimenter prior to training to provide a baseline and after criterion was reached to provide maintenance data. Procedures for these sessions were similar to those for instructional sessions except that they were conducted individually and without error correction. Probe sessions assessed target and nontarget vocabulary.

Generalization. All generalization measures were taken after students reached criterion on target definitions. Prior to instructional sessions, it was established that all target and nontarget definitions were unknown by all participants and data were collected for comparison purposes describing the percentage of target and nontarget words correctly read out loud.

Incidental learning. Incidental measures were taken after students reached criterion on target definitions. School staff conducted probe sessions following the procedures described above in which they asked students to define the nontarget words they had taught to their partners but had not studied themselves.

Maintenance. Maintenance checks were conducted approximately 4 and 10 weeks after completion of instructional sessions for each pair. An additional maintenance check was conducted for all participants during the last 2 weeks of the scnool year. The duration of time between reaching criterion and completing the final maintenance session ranged from 19 weeks for the last pair to 39 weeks for the first pair. Maintenance checks were conducted by the school staff members who conducted probe sessions during the instructional phase of the study, and in one case by the experimenter. Procedures for these sessions were identical to probe sessions described above.

Procedural integrity. The majority of instructional and probe sessions were audio recorded (four instructional sessions were not recorded due to equipment malfunction). The researcher served as the reliability observer and randomly selected tapes of 20% of each pair's instructional sessions and 100% of probe sessions. Behaviors selected for observation were: (1) accurate marking of correct and incorrect definitions, (2) waiting between 4 and 6 seconds after word presentation to begin correction procedures, and (3) providing verbal confirmation

for accurate responses. The number of correct behaviors was divided by the number of opportunities to exhibit correct behaviors and the result was multiplied by 100 in order to determine the percentage of behaviors performed correctly. Integrity checks ranged from 92% to 100% accuracy with an average of 94%, well in excess of Horner and colleagues (2005) suggested minimum of 80% accuracy for determining evidence-based practices in single-subject special education research. The most common incorrect behavior was waiting longer than 6 seconds to begin correction procedures. Interscorer agreement for accurate marking of correct and incorrect definitions was 97% for instructional sessions and 99% for probe sessions.

RESULTS AND DISCUSSION

Results

Correct Target Definitions

Results for the percentage of correct target definitions during instructional and probe sessions are shown in Figure 1. Each student responded with 0% accuracy during three or four baseline trials and, with the exception of Keisha, went on to reach 100% accuracy during the course of the study. Keisha's participation in the study was terminated by school staff due to repeated severe disruptive behavior during instructional sessions.

Tamika and Keisha were randomly assigned to receive intervention first. They both showed a sharp increase in correct responses upon intervention introduction. Tamika had a dip on her second session but then went on to make consistent increases until she reached 100% on her 12th trial. Because she dropped slightly below 100% two trials later, she required an additional four trials to reach criterion. Keisha's progress was much more variable with several large positive and negative swings. She reached 100% on her 13th trial and maintained it on her 14th. Following that trial she dropped lower than her first day of intervention and staff reported that her inappropriate behavior escalated. She completed four more trials with varying levels of success before staff terminated her participation.

Jamila and Jatara were randomly selected to receive intervention second. After both scored 0% on two additional baseline trials, Jamila responded with over 50% accuracy on her first instructional trial and reached 100% on her fourth. She completed nine more trials, eight at

100% accuracy. (This was a slight deviation from established methods because she should have stopped her trials after four at 100% but she asked to be allowed to continue and school staff made the judgment to permit this.) Jatara did not show a sharp increase on her first instructional session, but her progress was consistent and incremental and she reached 100% on her 11th trial.

Keon and Jamar received intervention third. Both scored 0% correct on the two baseline trials immediately prior to the intervention phase. Neither student made a sharp jump on the first instructional trial, but both did on the second trial. Keon made another large gain on the third trial then slowly increased to 100% accuracy on his seventh trial. Jamar's responses were more variable but he reached 100% on his 11th trial.

Definitions had a mean length of 11 words. Across all students except Keisha, mean total time per word learned was 3.64 minutes. This is roughly equivalent to the time per word reported by Schuster et al. (1990) in a teacher-delivered CTD vocabulary study that had younger students and shorter definitions.

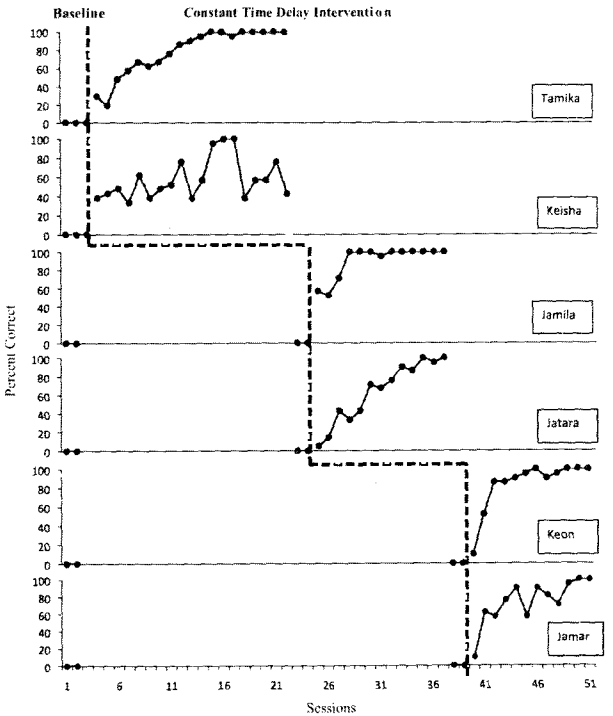


Figure 1. Percent Correct Definitions.

Generalization Measures

Word reading. Results for the percentage of correct word reading of science and social studies words are presented in Table 1. Pretest scores ranged from 14% to 79% correct (mean = 41.4%) and the five participants with posttest data all responded with 100% accuracy.

Word production when given definition. Results for the percentage of correct word production when provided the definition are presented in Table 1. All but Tamika scored 0% on the word production pretest. Presumably Tamika (who scored 29%) was guessing words she remembered hearing from the initial assessment to determine the words that had definitions unknown to all participants. The five students who completed this measure scored 100% at posttest.

Correct usage identification. Results for the five students with data for the number of correct answers on a multiple-choice test designed to measure identification of correct vocabulary (target and nontarget/incidental) usage are presented in Table 2. Maximum total correct score was 14, and results are divided into target and nontarget scores (maximum = 7). The range for total score was 6 to 11 (mean = 8.4). Four students accurately identified correct usage for five or six of the seven items; Jatara got three correct. Nontarget/incidental scores were lower and more variable: Jamar got two correct; Jatara, three; Keon and Tamika, four; and Jamila, five. All but one student scored higher on items assessing his or her own target words; Jatara got three correct from each set.

Incidental Definitions

Results for the pre- and posttest correct percentage of nontarget definitions are presented in Table 1. All students responded with 0% accuracy on the pretest, and the four students who were posttested

Table 1. Percentage of Correct Responses on Incidental and Generalization Measures

	Nontarget Definitions		*Word Reading*		*Word Production*	
Student	*Pretest*	*Posttest*	*Pretest*	*Posttest*	*Pretest*	*Posttest*
Tamika	0	48	79	100	29	100
Keisha	0	—	21	—	0	—
Jamila	0	95	57	100	0	100
Jatara	0	81	21	100	0	100
Keon	0	—	36	100	0	67
Jamar	0	90	14	100	0	93

Note. Dashes indicate that data were not obtained.

Table 2. Number Correct on Correct Usage Identification Measure

Student	*Target Score (of 7)*	*Incidental Score (of 7)*
Tamika	5	4
Keisha	—	—
Jamila	6	5
Jatara	3	3
Keon	5	4
Jamar	5	2

Note. Dashes indicate that data were not obtained.

demonstrated between 48% and 95% (mean = 78.5%) correct when asked for definitions of words they had never studied but had taught to their partners. No posttest data are available for Keisha who did not finish the study as explained above, nor for Keon, who finished instructional sessions and completed generalization measures but was removed from the school before completing the incidental session or maintenance checks.

Maintenance Data

Maintenance checks were conducted at approximately 4 and 10 weeks after each participant reached criterion and again at the end of the school year. Maintenance data are presented in Table 3. The time elapsed between each participant's final probe session and the last maintenance session ranged from 19 to 39 weeks for the four participants (Tamika, 39 weeks; Jamila and Jatara, 30 weeks; Jamar, 19 weeks) who completed the final check.

All four students maintained 95% to 100% accuracy at 4 weeks, and three of four scored between 90% and 100% at 10 weeks; Jatara had a

Table 3. Maintenance Session Percent Correct Target Definitions

	Number of Weeks from Final Session				
			End of Year		
Student	*4*	*10*	*19*	*30*	*39*
Tamika	95	90			67
Keisha	—	—			—
Jamila	100	100		57	
Jatara	100	71		43	
Keon	—	—	—		
Jamar	95	90	67		

Note. Dashes indicate that data were not obtained.

correct response rate of 71%. For the end-of-year maintenance check, Jamar and Tamika scored 67% correct, Jamila scored 57%, and Jatara maintained 43% correct responses.

Discussion

This study supports the idea that peer-mediated CTD interventions can support the goal of learning vocabulary and also aid in content area comprehension. Research generally supports the idea that preteaching key vocabulary can improve comprehension (e.g., Koury, 1996). Because vocabulary knowledge is a necessary prerequisite to understanding content area readings (although not sufficient in and of itself), future research should address the impact peer-mediated CTD vocabulary interventions may have on later reading comprehension.

Effectiveness Data

In addition to having academic deficiencies, participants in this study were nominated for inclusion due to disruptive classroom behaviors including fighting and general noncompliance, so it is perhaps not surprising that two of the six participants failed to complete all measures as a result of outcomes of their disruptive behaviors. Keisha evidenced definition learning with a positive (although highly variable) trendline, but staff reported her performance and behavior deteriorated substantially and terminated her participation before she reached criterion. Keon reached criterion and completed all measures except incidental learning and maintenance checks before being sent to a more restrictive environment for reasons unrelated to the study.

Middle school students with teacher-identified disruptive behavior demonstrated knowledge of previously unknown content area vocabulary words after participating in a peer-mediated CTD intervention. Words were relatively complicated and academically useful, and although there was considerable variation among students, all reached or came very close to reaching criterion. Given the necessary role vocabulary plays in reading comprehension, the reliance of secondary content area classes on textbooks and the academic difficulties experienced by students with disruptive behaviors, judicious use of instructional time is critical for them.

Incidental Learning

Four students completed the measure of incidental learning and all demonstrated mastery of a substantial number of previously unknown vocabulary definitions even though they received no direct instruction on them. Three students demonstrated greater than 80% mastery: Jamila (95%), Jatara (81%), and Jamar (90%).

Tamika demonstrated knowledge of slightly less than half (48%). One possible explanation for her comparatively poor performance on the incidental measure is that she did not finish the study due to behavioral concerns. Keisha's target word learning was highly variable and staff reported her behavior was largely erratic. Because Tamika would have been acquiring incidental words as a result of teaching them to Keisha, perhaps Keisha's behavior adversely impacted information Tamika was able to absorb. It is not possible to draw firm conclusions from one pair of students, but it seems plausible that Tamika's relatively lower incidental learning could be related to the fact that she had as a partner the only student who was removed from participation due to disruptive behavior during sessions.

Generalization

Participants were assessed on three generalization measures: word production upon definition exposure, word reading, and recognition of correct usage. The first two measures are similar to each other in that they address uninstructed association-level knowledge. Whether a student produces a definition in response to a word or a word in response to a definition, it is the association between the word and definition that is being assessed. Sight-word reading such as that which was assessed by the second measure is also association-level knowledge. For word production, scores increased from low levels (in all cases except Tamika who scored 29%, correct responses at pretest were 0%) to the much higher scores of 100% (Jamila, Jatara, and Tamika), 67% (Keon), and 93% (Jamar). Pretest scores on the word reading measure were variable, ranging from a low of 14% (Jamar) to a high of 79% (Tamika), but at posttest all students responded with 100% accuracy.

The measure assessing recognition of correct usage is distinct from the other generalization measures in the present study (i.e., word

reading and word production from definitions) because it deals with the deeper comprehension level knowledge whereas the other two address association-level knowledge. During instruction, students studied definitions but did not see target and nontarget words used in sentences, use the words to express themselves, or do anything with the definitional information other than memorize it and say it when prompted or prompt their partner to respond with verbal definitions for the other word set. This measure was developed and included to determine to what degree students would be able to apply newly acquired knowledge at a level deeper than it was taught.

There are two ways to consider these data. The first is by looking at total scores. A range of 6 to 11 (out of 14) may mean some participants were simply more skilled at the task itself because all participants were given this measure within days of reaching criterion on their target words, so they had similar levels of word knowledge at the outset of the study. Being skilled at this task would require the ability to transform known definitional information to an unfamiliar context.

A second way to consider these data is by examining the scores disaggregated by target and nontarget/incidental words. Doing so reveals that all but one student scored higher on the items that dealt with target rather than incidental words. Even though these students demonstrated high levels of incidental learning (95%, 81%, 48%, and 90% correct at posttest), they seemed less able to recognize when words, whose definitions they largely knew, were being used correctly.

For three students the difference in favor of target words was a single point, and for Jamar the difference was three points. While it is true that a single point still represents a significant proportion when the total score is seven, it is important to avoid extrapolating too much from such small numbers. Indeed, this is why we report scores on this measure alone as numbers rather than percentages.

Maintenance

Maintenance data for students completing the checks were strong at the 4- and 10-week measures (mean of all scores = 92.6%, range 71% to 100%) but end-of-year data were considerably weaker. The characteristics of the words themselves and goal of instruction perhaps explain this result. Target words were specifically selected because of their low incidence outside of the contexts of the particular science and social studies courses from which they were drawn.

Given the instructional goal was to teach these words only to the association level (as opposed to deeper and presumably longer-lasting levels of comprehension or generation), because they represented a specific barrier to understanding particular readings in these courses and were unlikely to be encountered frequently outside of class, it is not unexpected that without deeper instruction they were not maintained. Once intervention was concluded and students were no longer being exposed to definitional information on a regular basis, and once they had moved beyond those lessons in their courses, it is unlikely students would have encountered the words again with enough frequency to maintain learning. This would pose a problem in a vocabulary development program aimed to improve overall vocabulary performance of high-frequency words, but the intent of this intervention was to eliminate a specific barrier to comprehension of specific textbook content as quickly as possible. In order for learning to be maintained, information must be repeated over time and across contexts (e.g., Stahl & Fairbanks, 1986). However, for the purposes of this study, the instructional time it would have taken to maintain this learning at high levels for sustained periods would probably be better spent on other academic activities, perhaps vocabulary instruction on unfamiliar content area words from subsequent chapters.

We are not suggesting targeted vocabulary instruction should be viewed as "cramming" for a test. Rather we suggest that educators be judicious in allotting limited and valuable teaching time, especially to students with special needs or those students with disruptive behavior and to struggling readers. Instructional methods must match instructional goals, and if the word "secede" will help a student understand his social studies reading assignment but will most likely not present itself again for some years, there are better uses for the teacher's and students' time than it would take for the student to maintain that single definition.

One of the primary benefits of CTD interventions is that they can be quick and produce near-errorless learning (Wolery et al., 1992). Using CTD interventions in a peer-mediated format shows promise for improving the amount of information acquired in a set amount of time. In this study participants were able to implement the intervention with fidelity, master previously unknown definitions after brief instructional sessions, and acquire more vocabulary incidentally.

Constant time delay interventions are relatively simple to implement and do not require the same amount of teacher training and

planning time as interventions targeting more complicated skills, and in this study students were able to learn procedures quickly. Procedural integrity was robust and the intervention was successfully supervised in large part by school staff without direct researcher involvement, supporting the idea that peer-mediated CTD interventions can be successfully implemented in classrooms without substantial investments of training time.

Limitations

An important limitation of this study is a high rate of attrition. As mentioned above, only four of six participants completed all measures. If one member of a pair was unable to self-regulate behavior or withdraws from participation for some reason, the other member is unlikely to reap the full benefit from the intervention. Therefore, care is warranted when creating CTD pairs. It is possible that an aspect of the intervention, such as the relatively high frequency of corrective feedback, was aversive to Keisha and contributed to her disruptive behavior. Future studies should collect data on student engagement as well as off-task behaviors of partners to guide teachers in possible implications for matching and addressing problem behaviors, as well as to determine the nature and extent of influence on tutor and tutee roles.

A second limitation is the lack of systematic social validity data. Anecdotally, students reported enjoying the intervention (one was overheard exclaiming to a nonparticipating peer that she was learning "college words") and school staff noted that the intervention was simple to implement and students typically appeared eager for sessions. No special materials were needed, and minimal teacher time would be necessary in the classroom if students make flash cards. Additionally, it is encouraging that educators were able to successfully train participants and accurately collect data. However, it is important that future studies in this area systematically collect acceptability and utility data from students and teachers because there certainly exists the possibility that negative impressions of the intervention went unreported.

Because this study was conducted in an urban school with a significantly at-risk student body, and because all participants in the study were of similar socioeconomic status, external validity (i.e., the ability to generalize findings to other settings or populations) may be somewhat limited. Future studies should address peer-mediated CTD vocabulary interventions with other student groups.

A final limitation of this study is the lack of data addressing intervention effect on academic performance. This is especially important because the primary academic purpose of preteaching low-incidence vocabulary is to remove one barrier to text comprehension. Results indicate that students were able to learn target and nontarget definitions, but whether and to what degree this learning resulted in improved comprehension is not known. Though it is fairly well established that preteaching vocabulary improves comprehension generally (e.g., Koury, 1996), future research should address whether association-level definition learning has similar effects.

Implications for Practice

Results of the present study combined with the body of research noted earlier converge into five core recommendations for practice.

1. Select useful words for instruction. Students learn most words from context (Baumann & Kameenui, 1991). Even with efficient instruction, the maximum amount of words that could be reasonably taught in school would make up a small proportion of students' total vocabularies. Words targeted for explicit instruction should fall into one of two categories: words essential for academic comprehension or words which are relatively high frequency and useful in multiple contexts.

2. Teach target words using methods appropriate to instructional goals. Approaches to vocabulary learning should be selected with the goals of instruction in mind. The goal of instruction in this study was to help learners become sufficiently familiar with new vocabulary that they would be able to get the main ideas and content from subject area texts. Thus, target words from this study fall into the first category above and for time considerations were taught to the association level. High-incidence vocabulary words fall into the second category and should be taught to the generation level to be most useful.

3. Teach vocabulary in a peer or group arrangement. Stahl and Fairbanks (1986) hypothesized that vocabulary interventions conducted in groups would be more effective than those conducted individually because group discussion could provide more opportunities for practice and anticipation of being called on in a group could increase motivation. Later research has generally supported this contention; for example, two studies that addressed incidental vocabulary learning (Gast, Wolery, Morris, Doyle, & Meyer, 1990; Shelton,

Gast, Wolery, & Winterling, 1991) support the idea that instructional grouping can be beneficial even without discussion of definitions and lend support to the notion that individual vocabulary instruction is not necessarily more intensive or effective than instruction provided to small groups of students.

Group arrangements also have the potential to address social deficits, making group interventions perhaps more functional than other interventions if they are structured appropriately. For example, students with EBD have high levels of negative interactions and substantial difficulties interacting prosocially with peers and teachers (Gresham, Cook, Crews, & Kern, 2004). Students in this study displayed teacher-identified disruptive behavior rather than having documented EBD, but it seems reasonable that carefully structured, explicit programs may encourage positive interactions (Hansen & Lignugaris/Kraft, 2005) and provide opportunities to practice critical school behaviors such as turn-taking and listening.

4. Use CDT effectively in a peer mediated format by providing students sufficient training in the CDT method. In this study, the explicit model provided to students before they engaged in tutoring was perhaps helpful in providing them guidance for their behavior, and it is possible (although untested) that the correct responses of their peers reinforced their "tutoring behaviors." The opportunity to acquire more vocabulary incidentally cannot be overlooked. When a peer tutor must look at a definition with key words highlighted and provide feedback to the peer tutee, the opportunity to enhance the tutor's vocabulary occurs quite naturally. This doubling up of opportunity should have a positive impact on the tutor's vocabulary.

5. Assign student pairs with care. The body of research pertaining to peer-mediated vocabulary instruction in inclusive classrooms is small but encouraging. All students in this study exhibited behavior that was challenging and substantially problematic for their classrooms, although four had no identified disabilities and two were identified with LD. Students with LD were not paired together for intervention in either this study or a similar one conducted by Hughes and Fredrick (2006). In both studies, students with and without LD implemented CTD procedures accurately and demonstrated rapid acquisition. Hughes and Fredrick's vocabulary task was a less demanding one than definition production, but students with and without disabilities in an entire class worked together appropriately, learned information efficiently, and maintained acceptable classroom behavior. While one student did have her participation in the

current study terminated for repeated disruptive behavior, the rest of the students maintained positive behavior and complied with staff directives during instruction. One option not studied may be to rotate students across pairs. This may work if the vocabulary being studied is consistent across pairs and the novelty of the partnership may be a motivating factor.

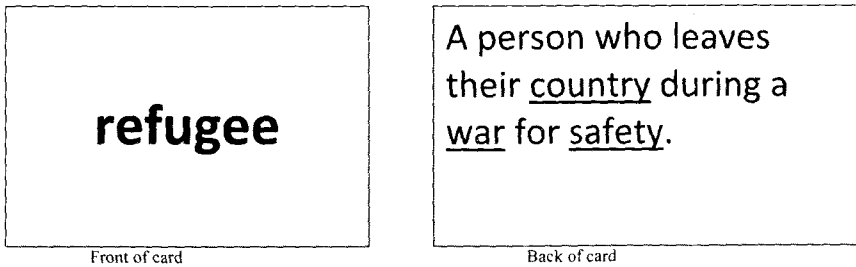
APPENDIX: IMPLEMENTATION GUIDELINES

Goal: To increase student knowledge of content area vocabulary definitions.

Objectives: Given the implementation of a peer-mediated Constant Time Delay instructional procedure, students will demonstrate 100% mastery of specified vocabulary words.

Preparatory Procedures:

1. Identify words from an upcoming reading that are likely to be unknown by either target students (specific) or a majority of students in the class (whole class).
2. Select five to seven words that are most fundamental for comprehension of the reading.
3. Devise simple definitions for selected words.
4. Create flashcards (to save class time, recommended for target students), or if desired, have students create flashcards (to save your time, recommended for whole class implementation).
   a. Write each selected vocabulary word on one side of a note card.
   b. Write the corresponding definition on the opposite side of each note card (Figure 2).



Front of card Back of card

Figure 2. Front and back of sample flashcard.

5. Assign partners and designate one student as "A" and one as "B" (or designation of your choice).

Procedures:

1. If using this for your whole class, it may be helpful to preteach procedures to one pair of students for demonstration.
2. Write the list of selected words on the board and read them out loud one at a time, and instruct students to repeat the pronunciation (**first session**).
3. Direct students to sit facing their partners and hand out one set of flashcards to each student (if they will keep them) or each pair (if you will keep them).
4. Direct students designated as A to read the first word and definition to their partners (**first session**).
5. Direct students designated as B to repeat the definition (**first session**).
6. Direct students designated as A to say "Yes," if their partners repeated the definition accurately, or to say the definition again and have their partner repeat the definition if an error was made (**first session**).
7. Repeat steps 4 to 6 for the remaining words in the set (**first session**).
8. Direct students to switch roles and repeat steps 4 to 7 (**first session**).
9. Direct students to switch roles (A's holding the flashcards) and read each word as they hold up the card, and then wait for up to 5 seconds for their partner (B) to begin saying the definition.
   a. If student B says the definition accurately, student A should say "Yes," or something similar.
   b. If student B says the definition inaccurately or says nothing after 5 seconds, student A states definition and student B repeats it.
10. Direct students to continue for a specified number of sessions, amount of time, or specified number of sets at 100% accuracy.
11. All sessions following the first session with the same set of words can omit steps 4 through 8.

REFERENCES

Baker, S. K., Simmons, D. C., & Kameenui, E. J. (1995). *Vocabulary acquisition: Synthesis of the research.* Eugene, OR: National Center to Improve the Tools of Educators.

Baumann, J. F., & Kameenui, E. J. (1991). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, J. M. Jenson, D. Lapp, & J. R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 604–632). New York: Macmillan.

Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly, 13*, 31–42.

Bulgren, J. A., Deshler, D. D., Schumaker, J. B., & Lenz, K. B. (2000). The use and effectiveness of analogical instruction in diverse secondary content classrooms. *Journal of Educational Psychology, 92*, 426–441.

Fuchs, L. S., Fuchs, D., Kazdan, S., & Allen, S. (1999). Effects of peer-assisted learning strategies in reading with and without training in elaborated help giving. *The Elementary School Journal, 99*, 201–219.

Gast, D. L., Wolery, M., Morris, L. L., Doyle, P. M., & Meyer, S. (1990). Teaching sight word reading in a group instructional arrangement using constant time delay. *Exceptionality, 1*, 81–96.

Greenwood, C. R., Arreaga-Mayer, C., Utley, C. A., Gavin, K. M., & Terry, B. J. (2001). Class wide peer tutoring learning management system: Applications with elementary-level English language learners. *Remedial and Special Education, 22*(1), 34–47.

Gresham, F. M., Cook, C. R., Crews, D. S., & Kern, L. (2004). Social skills training for children and youth with emotional and behavioral disorders: Validity considerations and future directions. *Behavioral Disorders, 30*, 32–46.

Hansen, S. D., & Lignugaris/Kraft, B. (2005). Effects of a dependent group contingency on the verbal interactions of middle school students with emotional disturbance. *Behavioral Disorders, 30*, 170–184.

Harmon, J. M., Hedrick, W. B., & Wood, K. D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading & Writing Quarterly, 21*, 261–280.

Heward, W. L. (2009). *Exceptional children: An introduction to special education.* Upper Saddle River, NJ: Merrill/Prentice Hall.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165–179.

Hughes, T. A., & Fredrick, L. D. (2006). Teaching vocabulary with students with learning disabilities using Classwide Peer Tutoring and constant time delay. *Journal of Behavioral Education, 15*, 1–23.

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*, 299–322.

Joshi, R. M. (2005). Vocabulary: A critical component of comprehension. *Reading and Writing Quarterly, 21*, 209–219.

Kameenui, E. J., Simmons, D. C., & Darch, C. B. (1987). LD children's comprehension of selected textual features: Effects of proximity of information. *Learning Disability Quarterly, 10*, 237–248.

Koury, K. A. (1996). The impact of preteaching science content vocabulary using integrated media for knowledge acquisition in a collaborative classroom. *Journal of Computing in Childhood Education, 7*, 179–197.

Kroeger, S. D., Burton, C., & Preston, C. (2009). Integrating evidence-based practices in middle science reading. *Teaching Exceptional Children, 41*, 3, 6–15.

Mastropieri, M., Scruggs, T. E., & Fulk, B. J. (1990). Teaching abstract vocabulary with the keyword method: Effects on recall and comprehension. *Journal of Learning Disabilities, 23*, 92–96.

Mastropieri, M. A., Scruggs, T. E., & Graetz, J. E. (2003). Reading comprehension instruction for secondary students: Challenges for struggling students and teachers. *Learning Disability Quarterly, 26*, 103–116.

McDonnell, J., Johnson, J. W., Polychronis, S., & Risen, T. (2002). Effects of embedded instruction on students with moderate disabilities enrolled in general education classes. *Education and Training in Mental Retardation & Developmental Disabilities, 37*, 363–377.

National Reading Panel (NRP). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* Bethesda, MD: National Institute of Child Health and Human Development.

Palincsar, A. S., & Herrenkohl, L. R. (2002). Designing collaborative learning contexts. *Theory into Practice, 41*(1), 26–32.

Ryan, J. B., Reid, R., & Epstein, M. H. (2004). Peer-mediated intervention studies on academic achievement for students with EBD: A review. *Remedial and Special Education, 25*, 330–341.

Schuster, J. W., Stevens, K. B., & Doak, P. K. (1990). Using constant time delay to teach word definitions. *Journal of Special Education, 24*, 306–318.

Scruggs, T. E., Mastropieri, M. A., Bakken, J. P., & Brigham, F. J. (1993). Reading versus doing: The relative effects of textbook-based and inquiry-oriented approaches to science learning in special education classrooms. *Journal of Special Education, 24*, 306–318.

Shelton, B. S., Gast, D. L., Wolery, M., & Winterling, V. (1991). The role of small group instruction in facilitating observational and incidental learning. *Language, Speech, and Hearing Services in Schools, 22*, 123–133.

Spencer, V. G., Scruggs, T. E., & Mastropieri, M. A. (2003). Content area learning in middle school social studies classrooms and students with emotional or behavioral disorders: A comparison of strategies. *Behavioral Disorders, 28*, 77–93.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72–110.

Tawney, J. W., & Gast, D. L. (1984). *Single-subject research in special education.* Columbus, OH: Merrill.

Wolery, M., Ault, J. M., & Doyle, P. M. (1992). Constant time delay procedure. In *Teaching students with moderate to severe disabilities: Use of response prompting strategies* (pp. 47–71). White Plains, NY: Longman Publishing Group.

Wolery, M., Holcombe, A., Cybriwsky, C., Doyle, P. M., Schuster, J. W., Ault, M. J., & Gast, D. L. (1992). Constant time delay with discrete responses: a review of effectiveness and demographic, procedural, and methodological parameters. *Research in Developmental Disabilities, 13*, 239–266.

Rachel Wannarka is a clinical assistant professor of education at the University of Michigan. She completed her doctorate in special education at The Pennsylvania State University (2009). Her current research centers around literacy for struggling adolescent readers and early-career teacher support in urban settings.

Kathy Ruhl is a professor of special education at The Pennsylvania State University. She completed her doctorate at the University of Florida (1983). Current research interests are in the areas of interventions for learners with mild disabilities and teacher preparation.

Richard Kubina is a professor in the special education program at The Pennsylvania State University. He completed his doctorate at The Ohio State University under John Cooper. His current research interests include measurably effective curricula and the effects of behavioral fluency within educational and social settings.