

Teacher Education and Special Education
2004, Volume 27, No. 4, 000–000

Preparing Preservice Teachers to Implement Class Wide Peer Tutoring

Larry Maheady, Gregory F. Harper, Barbara Mallette, & Melinda Karnes

Abstract: *This study focused on preservice general education teachers who were prepared to use an evidence-based teaching practice and the effects the practice had on their pupils' academic performance. Participants learned to use Juniper Gardens Children's Project's Class Wide Peer Tutoring (CWPT) program through a two-hour workshop and with in class assistance. The amount of time required for each preservice teacher to reach a pre-established training criterion (i.e., unassisted use of CWPT for three consecutive sessions with fidelity ratings of 85% and above) was calculated and specific implementation comments and concerns were recorded. Academic effects on pupils' spelling test performance were assessed using weekly pretests and posttests and social validity data were collected from all primary consumers. Key outcomes were that (a) preservice teachers were able implement CWPT with a high degree of accuracy with about 60 minutes of in class assistance, (b) use of CWPT resulted in high spelling grades on weekly posttests for all pupils, (c) preservice and cooperating teachers and their pupils reported favorable "treatment acceptability" for CWPT, and (d) some preservice teachers made procedural adaptations that appeared to be related to lower levels of pupil satisfaction. Findings are discussed in light of recent movements in the use of evidence-based teaching practices, professional accountability, and preservice teacher preparation.*

Teacher educators have been criticized for their failure to provide empirical evidence to show that (1) their programs make a meaningful difference in preservice educators' instructional practice and (2) their preservice teachers, in turn, make a noticeable impact on pupil achievement (e.g., Coalition for Evidence Based Educational Policy, 2002). A recent analysis of the empirical database on teacher education (Wilson, Floden, & Ferrini-Murphy, 2002) is instructive in this regard. These researchers were asked by the Office of Educational Research and Improvement and the U.S. Department of Education to conduct a systematic review of *high quality* research concerning five important teacher preparation questions. Wilson et al. (2002) reported that very few, if any, research studies could be found that *directly* addressed policy makers' questions. They found few definitive conclusions, for exam-

ple, regarding the impact of subject knowledge, pedagogical preparation, and/or student teaching experiences on teacher candidates' instructional competence. Similarly, little research had *directly* assessed the impact of specific educational policies (e.g., setting limits on number of education courses, requiring program accreditation, and changing duration of preparation from four to five years) and/or alternative certification programs on the quality of preservice teacher preparation.

Wilson et al. (2002) noted further that conducting a search for empirical research that met stringent standards gave them a sharper sense of the gaps that exist between teacher educators' claims and the evidence to support such assertions. In commenting on such findings, Shulman (2002) suggested that educational researchers may have made a serious error in judgment when they lost

Class Wide Peer Tutoring

Maheady, Harper, Mallette, & Karnes

sight of the importance of linking the distinctive features of teaching to the quality of student learning. He noted, for example, that educational researchers may have become so enamored of teachers' cognitive processes that they ignored student learning.

Here, we discuss briefly an evidence-based teaching practice, Class Wide Peer Tutoring (Greenwood, Delquadri, & Carta, 1997) and describe a study in which we prepared a group of preservice general educators to use it and then examined the effects of their instruction on pupils' weekly spelling test performance. We offer this study as a reasoned attempt to address some of the criticisms being directed at special education and its personnel preparation and research efforts.

What is Class Wide Peer Tutoring?

Class Wide Peer Tutoring (CWPT) was developed by researchers at the Juniper Gardens Children's Project in Kansas City, Kansas, in the early 1980s (Delquadri, Greenwood, Stretton, & Hall, 1983). It is currently available in manual form under the commercial title of *Together We Can* (Greenwood et al., 1997). CWPT was designed initially to improve the basic skills (i.e., reading, math, and spelling) of lower performing elementary students (Delquadri et al., 1983). It has been subsequently extended to secondary content-area courses in both general and special education settings (e.g., Bell, Young, Blair, & Nelson, 1990; Maheady, Harper, & Sacca, 1988; Maheady, Sacca, & Harper, 1988; Wright, Cavanaugh, Sainato, & Howard, 1995).

To what extent does CWPT represent an evidence-based teaching practice? Although no widely established criteria have been adopted in this regard, some data suggest that CWPT may, in fact, qualify for this designation. First, an Educational Resources Information Center (ERIC) search conducted by Greenwood, Arreaga-Mayer, Utley, Gavin, and Terry (2001) found at least 25 published studies that have shown CWPT to be superior to conventional forms of teacher-led instruction in improving pupils' academic outcomes. Moreover, most of these investigations were high quality between-

group and/or single-subject research designs (Greenwood, Maheady, & Delquadri, 2002). Perhaps the most compelling support, however, came from a 12-year experimental longitudinal study (Greenwood, Delquadri, & Hall, 1989; Greenwood & Delquadri, 1995). Here researchers compared groups of students at-risk and not at-risk who had or had not received CWPT instruction. They found that CWPT (a) increased students' active engagement during instruction in grades 1 to 3, (b) improved pupil achievement at grades 2, 3, 4, and 6, (c) reduced the number of CWPT students in need of special education services by 7th grade, and (d) decreased the number of students who dropped out of school by the end of 11th grade. More recently, Horner, Carr, Halle, McGee, Odom, and Wolery (in press) proposed that a particular practice can be considered evidence-based if a minimum of five high quality single subject studies have been published in peer-reviewed journals by at least three different researchers across three different geographical locations. Moreover, these studies must include a total of at least 20 participants. The literature indicates clearly that CWPT meets such criteria. Finally, in a recent issue of the Council for Exceptional Children's Division of Research publication, *Alert*, CWPT was given the *Go With It* designation to reflect its substantive data base (Maheady, Harper, & Mallette, 2003). While there is good support for CWPT's effects, studies focused on professional development and relations between "training" and outcomes have been less evident.

Method

Three questions were addressed:

1. To what extent can we document a procedure for preparing preservice general educators to use CWPT in an accurate and durable manner?
2. To what extent does preservice educators' use of CWPT effect pupils' weekly spelling test performance?
3. To what extent are CWPT goals, procedures, and outcomes socially acceptable for primary consumers (e.g., pupils, preservice, and cooperating teachers)?

Each focused on a key component of the CWPT knowledge base and provided strong

TESE, Volume 27, No. 4
Fall 2004

implications for professional development in general and special teacher education programs.

Participants and Settings

Ten female preservice general educators and their cooperating teachers from two, small school districts ($N = 2,300$) in Western New York served as participants. Preservice teachers were enrolled in the final semester of their undergraduate program and carried out the study in one of their eight-week student teaching placements. Preservice educators were selected randomly from a pool of all student teachers placed in grades 2 through 4 within the two local districts. Once selected for participation, both preservice and cooperating teachers were given a brief explanation of the study and were asked to participate. If teachers said yes, then plans were made to visit their classrooms the following weeks. If either preservice or cooperating teachers did not want to participate, then other individuals would have been selected. In all instances, both groups of teachers were told explicitly that participation was strictly voluntary and that preservice teachers' grades would not be affected by their involvement. All selected individuals agreed to participate. Two cooperating teachers (Classrooms C and J) had prior experience with CWPT.

Two hundred and seven second through fourth grade students (52% male, 48% female) also participated. Ninety-two percent of the pupils were Caucasian, 4% African-American, 3% Hispanic, and the remaining 1% Asian Pacific. Students ranged in age from 7 years 3 months to 11 years 7 months ($M = 9$). Twenty-eight students had been identified previously as learning disabled according to district eligibility criteria. All CWPT sessions took place in general education classrooms during regularly scheduled spelling lessons and were carried out by preservice teachers.

Class Wide Peer Tutoring and Preparation Experiences

CWPT is an instructional intervention in which pupils work in pairs on a class-wide basis. The program includes four major com-

ponents: (a) weekly competing teams, (b) structured tutoring procedures (i.e., systematic task presentation, immediate positive and corrective feedback, contingent point earning, and role reversals), (c) daily point earning and public performance displays, and (d) direct practice of functional academic skills (i.e., spelling words orally and in written form) (for a more detailed description of procedures see Greenwood et al., 1997; Greenwood et al., 2002). Here, CWPT was implemented only in spelling and was used 3 or 4 days per week for approximately 20 minutes per day.

Preservice teachers were prepared to use CWPT initially via a 2-hr workshop held on the college campus. They were given a procedural implementation manual containing relevant theory underlying the use of the tutoring system and explicit instruction in CWPT use. Preservice teachers were shown 15-minute video clips depicting CWPT in spelling, and then they participated in brief role-play situations. The first and third authors served as trainers and provided students with positive and corrective feedback on their implementation efforts. Additional assistance was then provided directly in each preservice teacher's student teaching placement by the authors and a graduate assistant. In-class assistance consisted primarily of consultant feedback regarding preservice teacher performance and occasional modeling of "correct" tutoring procedures. In-class assistance continued until preservice teachers reached a pre-established training criterion (i.e., implementing CWPT for three consecutive sessions without consultant assistance and obtained fidelity ratings of 85% and above). Two preservice teachers failed to reach the criterion because they adapted CWPT at cooperating teachers' requests. In these instances, assistance was provided until preservice teachers reached a criterion of 80% for three consecutive sessions. Inter-rater reliability was conducted by having the first and third authors simultaneously and independently assess fidelity on at least one occasion for each preservice teacher during both initial training and while using CWPT in the class.

During initial preparation, preservice teachers were asked to follow CWPT pro-

Class Wide Peer Tutoring

Maheady, Harper, Mallette, & Karnes

cedures as closely as possible but to respond to any cooperating teachers' concerns. In particular, they were asked to eliminate and/or adapt any CWPT procedures that cooperating teachers did not like or did not want them to use.

Dependent Measures

Preparation outcomes were assessed along three dimensions: (a) extent and fidelity of CWPT implementation, (b) academic effects on students' weekly spelling test performance, and (c) consumer satisfaction with intervention goals, procedures, and outcomes.

CWPT Implementation Measures

The following measures of CWPT implementation were derived. First, we calculated the percentage of preservice teachers reaching the pre-established training criterion. CWPT implementation was determined through direct observation. Immediately following each classroom visit, consultants (i.e., first and third author and a graduate assistant) completed a one page, Consultant Reaction Form. This experimenter-developed form included information regarding (a) the length (time in, time out) and nature of contacts (i.e., what consultants and preservice teachers did with regard to CWPT implementation), (b) the level of assistance provided (i.e., no assistance to full assistance), (c) any teacher-initiated questions or concerns regarding implementation, and (d) any teacher-initiated supportive comments about CWPT. Information from these forms was used to calculate initial training times to criterion and to quantify aspects of consultative contacts.

The second implementation measure was intervention fidelity. Using 3 categories, a 36-item CWPT Procedural Checklist developed by Carta, Greenwood, Dinwiddie, Kohler, and Delquadri (1987), in class consultants scored each item as either present or absent. Categories included CWPT materials (7 items), teacher procedures (15 items), and students' tutoring behavior (14 items). Consultants also rated preservice teachers on both dependent measures immediately after each class visitation. Fidelity assessments were

conducted five times on each preservice teacher throughout the investigation and inter-rater reliability data were calculated on at least one of these occasions. Procedural checklists yielded a percentage of components implemented using the formula: number of components present divided by the number present and absent $\times 100\%$. If preservice teachers failed to implement particular procedural components at the request of the cooperating teacher, they were scored as absent.

The final CWPT implementation measures were the number of training sessions and amount of consultant time required for preservice teachers to reach pre-established training criteria. Since they were required to use CWPT unassisted for three consecutive trials (but consultants were still required to be present), both measures were inflated.

Academic Outcome Measures

CWPT's academic effects were assessed using weekly spelling pretests and posttests. The primary dependent measures were percentage of words spelled correctly and the amount of normalized gain made above pretest performance. Pretests and posttests consisted of 20 words that were administered orally by preservice teachers. Pretests for the following weeks were given immediately after each posttest in the same manner. Inter-rater reliability was calculated on 25% of all spelling assessments by having the first and third author independently score tests. Reliability was calculated as the number of agreements divided by number of agreements plus disagreements times 100%.

Social Validity Measures

Consumer satisfaction was assessed across pupils, preservice and cooperating teachers. Pupils were asked to complete independently and anonymously a 14-item, Peer Tutoring Evaluation Inventory following the final CWPT sessions. Questions were read orally by preservice teachers and pupils marked their individual response sheets. The first four items asked students to note (on a 3-point Likert-type scale) their *general level of satisfaction* with CWPT. Questions included: Is CWPT something that should be done in school? How much did you like the things

TESE, Volume 27, No. 4
Fall 2004

you did during CWPT? The next six items requested students to rate the *acceptability of CWPT procedures* (i.e., being on a team, earning points, seeing points on the scoreboard), while the remaining four items asked them to assess academic and social *outcomes* associated with CWPT use (e.g., Do you think other students were more friendly to you after CWPT? Do other students think that you are smarter now than before CWPT?). The final item asked pupils to provide an overall assessment of CWPT (i.e., 1 = like it a lot, 2 = it's OK, and 3 = don't like it).

Preservice and cooperating teachers also completed 16-item consumer satisfaction surveys independently and anonymously at the end of CWPT sessions. Three items requested information regarding the importance of intervention goals, eight questions pertained to procedural acceptability, and the final five items assessed satisfaction with intervention outcomes. All teacher responses were recorded on a 3-point Likert-type scale.

General Design and Procedures

The present investigation used two research methods. At the teacher level, we conducted a descriptive study of implementation practices following a training-to-criterion design. Using a standard workshop model with in class assistance, we attempted to bring 10 preservice general education teachers to a pre-established training criterion. We also directly observed their CWPT implementation during an eight-week student teaching experience. Primary dependent measures were the accuracy with which they used CWPT procedures, the percentage of preservice teachers who reached the training criterion, and the amount of time it took them to do so.

At the pupil level, we assessed the academic outcomes associated with preservice teachers' use of CWPT methods. Students were given weekly pretests and posttests containing 20 "new" spelling words each week. Spelling instruction typically took place for about 20 minutes per day using Riverside Spelling (Wallace, Taylor, Fay, Kucera, & Gonzalez, 1988) in one district and MacMillan Spelling (Smith, 1987) in the other. Similar instructional procedures were fol-

lowed in each classroom. On the first Friday, pupils were pretested on the upcoming list of 20 spelling words. The following Monday through Thursday, they participated in CWPT by using those 20 words. On Friday, spelling tests were administered wherein preservice teachers dictated each word, used it in a sentence, and pronounced it again. Pupils were required to write the correct spelling for each dictated word. Immediately after the posttest, pupils were given a pretest on the subsequent list of 20 spelling words. Similar instructional procedures were followed each week.

CWPT sessions were conducted as follows. Each Monday, preservice teachers introduced new spelling words for the week. Each class read the words in unison two to three times and the preservice teacher provided each word meaning. Two competing teams were formed in each classroom by having students select either a red or blue piece of paper from a covered box. Team membership (i.e., red vs. blue team) stayed the same for the rest of the week. Teams competed for the higher point totals in spelling. Following team selection, preservice teachers randomly paired students within each team to make up tutoring pairs. Pairs worked together throughout the week on their 20-word spelling lists. Each tutoring session began with the preservice teacher setting a kitchen timer for 10 minutes. One student served as tutor while the other was the tutee. Tutors dictated each word to their respective tutees who, in turn, were required simultaneously to write and orally spell each word. Tutors said, "that's right" or "correct" and provided two points for each correct spelling. However, if tutees misspelled the dictated words, tutors said "wrong" and orally spelled the word correctly. Tutees were then required to write the misspelled word correctly three times. One point was awarded if the spelling error was corrected. No points were awarded if tutees failed to correct their errors. Pupils continue through their spelling lists as many times as possible before the 10-minute time period elapsed. When the bell rang, tutor pairs reversed roles for the same amount of time. Timers were reset and tutor pairs continued working as quickly as possible, following the same tutoring procedures.

Class Wide Peer Tutoring

Maheady, Harper, Mallette, & Karnes

Bonus points were awarded daily to tutors who displayed "good" tutoring behaviors. Preservice teachers moved about the classroom and awarded points, up to five per student, to tutors who: (a) clearly and accurately pronounced words, (b) appropriately used error correction procedures, and (c) contingently delivered points. Tutees also earned bonus points for writing spelling words legibly and for correcting errors without complaint and/or hesitation. Following the second 10-minute session, students totaled their points, including bonus points, and posted their respective scores on a laminated chart in the front of class. Cumulative team totals were calculated daily. CWPT procedures remained in effect from Monday through Thursday. On Friday, the weekly posttest was administered using the standard format of preservice teachers pronouncing each word, using it in a sentence, and pronouncing it a final time. However, pupils were informed that they could earn five points for themselves and their team for each word spelled correctly on the test. Whatever grade was earned on the Friday test was then entered into the teachers' grade-books.

Following posttesting and subsequent pretesting of upcoming words, team totals were calculated and "Team of the Week" certificates, signed by all team members, were given to winning teams. Certificates were then posted on walls in the hall or displayed prominently within the classroom. Each week new competing teams were formed, thereby increasing the probability that each student would participate on a winning team, while simultaneously minimizing the possibility that cliques or firmly entrenched teams would develop.

Results

Findings are presented in four sections: (a) inter-rater reliability on independent and dependent variables, (b) training and implementation outcomes, (c) academic effects on students' spelling test performance, and (d) social validity data.

Inter-rater reliability

Inter-rater reliability was calculated on preservice teachers' fidelity of CWPT use

during initial training and in student teaching placements. Inter-rater agreement levels averaged .96 and .92 respectively with a range of .88 to 1.00 across both preservice educators and settings. Inter-rater reliability data on pupils' spelling test performance averaged .98 (range .94–1.00).

Training Implementation Outcomes

Can we document a procedure for preparing preservice general educators to use CWPT in an accurate and durable manner? The answer appears to be yes. Given a preparation manual, a two hour workshop, in class assistance, and a pre-determined training criterion, all preservice teachers learned to use CWPT with high degrees of accuracy ($M = 87.7\%$; range = 82–96%). This suggests that they regularly performed about 32 of 36 CWPT procedures during typical spelling lessons. Given that fidelity checks were conducted throughout the student teaching experiences and no trends in accuracy were apparent, the findings suggest durable change in preservice teachers' instructional practice.

The data also indicate that teacher preparation was done in a rather efficient manner. For example, the average amount of in class assistance time required to bring preservice teachers to criterion was just over two hours ($M = 127$ min). Given that three visits (i.e., 60–90 min) involved only consultant observation and no assistance, the data suggest that preservice educators learned to use CWPT with less than 1-hr of in class help. It is also significant to note the large range in individual training time (i.e., 65–153 min). In two classrooms where both cooperating teachers had previous experience with CWPT, preservice teachers' time to criterion was considerably less than in other settings.

Eight of 10 preservice teachers reached the pre-established training criterion (i.e., unassisted use of CWPT with 85% fidelity for three consecutive sessions) within approximately 6 class visits ($M = 5.3$). The 2 preservice educators who failed to reach criterion adapted CWPT procedures at their cooperating teachers' request but maintained over 80% integrity in overall use. Analyses of individual reliability protocols indicated fur-

TESE, Volume 27, No. 4
 Fall 2004

Table 1. Weekly Pretest and Posttest Averages and Normalized Gain Scores across 10 Classrooms

Preservice Teachers	Grade	Pupils	M Pretest	M Posttest	% of Normalized Gains
A	4	23	59.6	85	59.3
B	2	20	86.9	96.8	89.5
C	3	19	51.9	94.7	51.5
D	4	24	67.2	87.5	67
E	2	20	89.4	98.8	89.3
F	2	19	83.8	97	83.7
G	3	18	53.6	95.1	53.2
H	4	21	64.5	92.8	64.2
I	3	23	80.6	94.2	80.5
J	4	20	58.2	92.4	57.9
Totals		207			
M			69.6	93.5	69.4
Range			(52–89.4)	(85–98.8)	(51.5–89.5)
SD			14.32	4.29	

ther that preservice teachers had omitted and/or adapted the following CWPT components: (a) daily and/or bonus points, (b) public posting of students' scores, and (c) formal team recognition.

Finally, additional data derived from Consultant Reaction Forms yielded 86 teacher-initiated concerns/supportive comments. Teachers offered 47 supportive comments (55%) regarding CWPT's effectiveness and acceptability, and 39 comments regarding procedural and management concerns. Teachers noted, in particular, that students occasionally became loud and did not follow prescribed tutoring procedures.

Academic outcomes

What effect does preservice educators' use of CWPT have on pupils' weekly spelling test performance? Mean pretest and posttest spelling grades and normalized gain scores are presented below. Pupils' weekly spelling test performance increased from pretest averages of 69% to posttest averages of 93.5% across the 10 classrooms (see Table 1). This reflects approximately a 25% increase in pupils' actual spelling averages and 2 to 3 grade level improvements (i.e., D+ to A). On regularly administered spelling tests following CWPT, 78% of all students earned A grades, while the failure rate (i.e., below 60%) was less than 1%. In fact, of the 1,028 spelling tests administered following CWPT instruction, only 8 resulted in failing grades.

Given that four classrooms (B, E, F, & I) had high pretest scores, we used normalized gain scores to better represent the amount of improvement that pupils made. Normalized gain scores are defined as the ratio of the actual gain to the maximum gain possible (Hake, 1998). Normalized gains were calculated as the posttest minus pretest divided by 100 minus pretest \times 100%. The average normalized gain score was 69.4% (range = 51.5–89.5%).

Social validity outcomes

How socially acceptable are CWPT goals, procedures, and outcomes for primary consumers? Since very few differences existed between preservice and cooperating teachers' ratings, their evaluations were combined. Both groups reported that it was very important for pupils to learn to work cooperatively and to do well on Friday spelling tests. Their responses were less positive about the importance of pupils monitoring and correcting each other's work. Preservice and cooperating teachers also reportedly liked most CWPT procedures, particularly giving points for improved social interactions, awarding recognition certificates, using the instructional procedures, and placing students onto weekly teams. They provided lower ratings for the use of the scoreboard and the public display of students' scores. Preservice and cooperating teachers reported that CWPT improved their pupils' spelling

Class Wide Peer Tutoring

Maheady, Harper, Mallette, & Karnes

test performance and to a lesser extent enhanced social relationships among students. Overall, 18 of 20 teachers (i.e., 9 preservice and 9 cooperating) said that they would definitely use CWPT again, while two others were less certain.

Pupil satisfaction was equally positive but more varied. Pupils' overall assessments of CWPT were quite favorable. Mean ratings across all 10 classrooms were below 2 indicating high pupil satisfaction. Interestingly, the most positive student evaluations occurred in classrooms where procedural fidelity ratings were the highest (i.e., Classrooms E, F, C, and J), while satisfaction ratings were least favorable in classrooms where substantial procedural adaptations occurred (i.e., Classrooms I, G, D, and H). Overall, only 15 out of 207 pupils (7%) reported that they did not like CWPT. Thirteen of these pupils were enrolled in classrooms where procedural changes were made.

Overall, 92% of all pupils reported that they would use CWPT if they were teachers. A comparable number said that they liked the things that they did during CWPT sessions. The least favorable ratings on these items came from pupils enrolled in Classrooms D, G, and I. Regarding specific procedural components, pupils clearly rated earning points, being on a team, seeing their scores posted, and giving and/or receiving help as the most favorable aspects of CWPT. In contrast, there was a general consensus that the error correction procedure was the least liked CWPT component. Most pupils (86%) reported that others thought they were smarter after their involvement with CWPT, and almost half of the pupils noted that they got along better with other students. Again, less favorable responses were noted among pupils enrolled in classes where substantial procedural adaptations were made.

Discussion

Clearly, preservice general educators can learn to use an evidence-based teaching practice like CWPT in "real life" settings with a high degree of accuracy and a minimum of external assistance. Importantly, preservice teachers' use of CWPT produced noticeable

improvements in pupils' weekly spelling test performance. This brings a measure of effectiveness to their implementation efforts and appears to meet policy makers' expectations that we (teacher educators) demonstrate that our preservice teachers make an impact on pupils' academic performance (e.g., Coalition for Evidence Based Educational Policy, 2002; Wilson et al., 2002). When preservice teachers used CWPT, over three-fourths of their pupils earned A grades on Friday spelling tests, while less than one percent failed. Such positive academic outcomes replicate results from prior *experimental* studies using CWPT in spelling (e.g., Greenwood et al., 1989; Harper, Mallette, & Moore, 1991; Maheady & Harper, 1987; Mallette, Harper, Maheady, & Dempsy, 1991; Sideridis, Utley, Greenwood, Dawson, Delquadri, & Palmer, 1997). Finally, consumer satisfaction data indicated that preservice and cooperating teachers, as well as their respective pupils found CWPT to be socially acceptable. Again, these findings are highly consistent with previous CWPT reviews (see for example, Greenwood et al., 2002; Maheady, Harper, & Mallette, 2001; Utley, Mortweet, & Greenwood, 1997).

This study has important implications for preservice teachers, teacher educators, and educational researchers alike. Preservice teachers, general and special education, would benefit greatly from direct preparation in socially acceptable, evidence-based teaching practices. Acquiring instructional competence with some powerful teaching practices may allow them to experience success early in their careers, thereby permitting them to prevent rather than remediate learning failures. Gaining instructional competence in the use of evidence-based teaching practices may also instill self-confidence in future teachers, prompt them to seek out other data based instructional strategies, and/or stimulate them to generalize the use of existing practices to new students, settings, or subject areas. Our future teachers may also become more attractive to public school employers who must show that they, too, use evidence-based teaching practices. Perhaps most importantly, preservice teachers would benefit greatly from knowing that a science of teaching and learning does exist, that all

TESE, Volume 27, No. 4
Fall 2004

instructional practices are not created equally, and that ultimately, the most effective teaching strategies will be those that maximize their own students' learning (Greenwood & Maheady, 2001). Preservice teachers must obviously learn much more than the procedural knowledge to use a few evidence-based teaching practices. Minimally, they must also acquire the conditional knowledge to know when and where to apply these newly acquired practices, as well as the evaluative competence to assess the subsequent impact of their teaching on important pupil outcomes.

This study also has important implications for those charged with teacher preparation in the 21st Century. Without question, teacher educators are currently "under the gun" to provide evidence that (1) they make a difference in their preservice teachers' instructional practice, and (2) their preservice teachers, in turn, make a noticeable impact on pupil performance. This study described a *process* for preparing preservice teachers to use a specific evidence-based teaching practice as well as *procedures* for collecting evidence to show that they did, in fact, do so. The use of a workshop format supplemented with in class assistance and ongoing procedural feedback served as a useful vehicle for helping preservice general educators learn to use CWPT with a high degree of accuracy. The use of a training-to-criterion design, consultant reaction forms, and procedural fidelity checklists allowed us to document that our preservice teachers did, in fact, adapt their instructional behavior. Obviously other preparation formats and evaluation strategies can be used to demonstrate that teacher educators do make a noticeable difference in preservice teachers' instructional practice. To date, for example, we (a) used *peer coaching* to improve preservice teachers' ability to implement Peer Assisted Learning Strategies (PALS) (Fuchs, Fuchs, Mathes, & Simmons, 1997) in an after-school tutoring program for students with special needs (Mallette, Maheady, & Harper, 1999), (b) implemented a similar *multi-component training package* to prepare eight inservice general educators to use CWPT in spelling (Maheady, Harper, Mallette, & Winstanley, 1991), and (c) required *teaching/*

learning assignments within our early field-based courses to prepare preservice educators to use other evidence-based teaching practices (Maheady, Harper, & Mallette, 2005). As the press for additional evidence increases, teacher educators would do well to develop a variety of effective and efficient preparation strategies and more robust measures of teachers' instructional competence.

This study offers some interesting implications for educational researchers as well. First, it provides a *new lens* for conducting educational research. In this study, teachers' instructional practice served as the critical *dependent* variable. Often, educational research treats what teachers *do* instructionally as the independent variable and what happens to pupils as the dependent measure. In this research line, what teachers do instructionally is the primary outcome variable and its nature and occurrence lends itself to empirical analysis. What researchers and teacher educators do to impact instructional practice become our independent variables. Clearly, we are in our infancy with regard to specifying the nature and types of instructional practices that we would like to foster. One might conjecture, however, that such strategies can be categorized into teacher-directed, peer-assisted, student-regulated, and/or technology-assisted approaches (Maheady, 1997).

This research line will also require much better "tools" for assessing important instructional outcomes. Here, we used Consultant Reaction Forms, CWPT Procedural Checklists, and consumer satisfaction surveys to examine preservice teachers' practice. More work must clearly be done on the technical adequacy of these measures and on the development and refinement of other direct implementation measures. Finally, educational researchers must begin to document *functional relationships* between their preparation strategies and teachers' instructional practice, as well as between teacher's practice and important pupil outcomes. To do so, they must increase their use of either randomized controlled trials or high quality single subject research designs.

The present study obviously has its own limitations. First, our outcome data were derived using descriptive rather than experimental designs. Therefore, we can only de-

Class Wide Peer Tutoring

Maheady, Harper, Mallette, & Karnes

scribe what happened as a result of our preparation efforts. We cannot ascertain which factors (e.g., manual, modeling, or in class assistance), for example, were “really” responsible for preservice educators’ ability to use CWPT, nor can we comment on the appropriateness and/or superiority of our direct teaching model over a different preparation approach. Similarly, our use of a pretest and posttest design, without random assignment or other procedural controls, does not permit us to conclude that CWPT was indeed responsible for pupil growth in spelling. Given the rather substantial empirical database that already exists in this regard, however, we are not overly concerned by our use of a less stringent research design.

The present investigation was further limited by the relatively small number of preservice general educators, a short implementation duration, and the use of a rather narrow facet of literacy development (i.e., spelling). Moreover, we focused on preservice teachers’ use of only one, highly structured teaching practice and we did not assess whether they either sustained or generalized their CWPT use beyond the investigation. Another limitation involved our use of volunteers. One is more likely to obtain favorable outcomes when they work with volunteers as opposed to those who may not be interested in a particular teaching method. Here, all preservice teachers volunteered to use CWPT after hearing a brief description of it. Whether they decided to do so because CWPT “made sense” to them, because they felt that their cooperating teachers and/or college instructors wanted them to do so, or because of some other reason is difficult to ascertain. In any event, we cannot assume that comparable outcomes will be derived with non-voluntary personnel.

In conclusion, as we reflected upon our efforts to document overt changes in preservice teachers’ instructional practice, we were reminded of some profound advice offered by Don Baer (1993). The late Professor Baer wrote:

In science research, there are always two opposing strategies to choose between: One is to attack a problem because it is important, no matter how weak the available research methods may be, be-

cause for important problems even ambiguity is better than ignorance. The other is to attack only those problems where powerful research methods can convey unambiguous answers, confessing ignorance with the remainder, because confessions of ignorance are better than ambiguous assertions, especially with important problems (p. 109).

Clearly, we believe that it is very important for teacher educators to demonstrate that they can and do make a noticeable difference in their preservice teachers’ instructional practice and that their teachers, in turn, do the same with their students. We know as well, however, that our existing research methods are not sufficiently rigorous or technically adequate to provide the unambiguous answers that we need in this regard. We offer our findings here, therefore, as an initial measure of ambiguity in a larger sea of ignorance.

References

- Baer, D. M. (1993). Advising as if for research productivity. *The Clinical Psychologist*, 46, 106–109.
- Bell, K., Young, K. R., Blair, M., & Nelson, R. (1990). Facilitating mainstreaming of students with behavior disorders using classwide peer tutoring. *School Psychology Review*, 19, 564–573.
- Carta, J. J., Greenwood, C. R., Dinwiddie, G., Kohler, F. W., & Delquadri, J. C. (1987). *The Juniper Gardens classwide peer tutoring programs for spelling, math, and reading*. Kansas City, KS: Juniper Gardens Children’s Project.
- Coalition for Evidence Based Educational Policy Report (2002). Rigorous evidence: The key to progress in education. <http://www.excelgov.org/displayContent.asp?Keyword=prppcEvidence>
- Delquadri, J. C., Greenwood, C. R., Stretton, K., & Hall, R. V. (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.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.
- Greenwood, C. R., Arreaga-Mayer, Utley, C. A., Gavin, K. M., & Terry, B. J. (2001). Class Wide Peer Tutoring Learning Management System. *Remedial and Special Education*, 22, 34–47.
- Greenwood, C. R., & Delquadri, J. (1995). Class Wide Peer Tutoring and the prevention of school failure. *Preventing School Failure*, 39(4), 21–25.
- Greenwood, C. R., Delquadri, J. C., & Carta, J.

TESE, Volume 27, No. 4
Fall 2004

- J. (1997). *Together we can!: Class Wide Peer Tutoring for basic academic skills*. Longmont, CO: Sopris West.
- Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology, 81*, 371–383.
- Greenwood, C. R., & Maheady, L. (2001). Are future teachers aware of the gap between research and practice? *Teacher Education and Special Education, 24*, 333–347.
- Greenwood, C. R., Maheady, L., & Delquadri, J. C. (2002). Class-Wide Peer Tutoring. In G. Stoner, M. R. Shinn, & H. Walker (Eds.) *Interventions for achievement and behavior problems* (2nd Ed.) (pp 611–649) Washington, DC: National Association of School Psychologists.
- Hake, R. R. (1998). Interactive engagement versus traditional methods: A six thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*, 64–74.
- Harper, G. F., Mallette, B., & Moore, J. (1991). Peer-mediated instruction: Teaching spelling to primary school children with mild disabilities. *Reading, Writing, and Learning Disabilities International, 7*, 137–151.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (in press). The use of single subject research to identify evidence-based practice in special education. *Exceptional Children*.
- Maheady, L. (1997). Preparing teachers to work with multiple ability groups. *Teacher Education and Special Education, 20*, 322–339.
- Maheady, L., & Harper, G. F. (1987). A classwide 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., Harper, G. F., & Mallette, B. (2003, Spring). Class Wide Peer Tutoring. *ALERT*, (No. 8), 1–6.
- Maheady, L., Harper, G. F., & Mallette, B. (2005). Developing, implementing, and maintaining a responsive educator program for preservice general education teachers. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 139–153). Upper Saddle River, NJ: Prentice Hall/Merrill.
- Maheady, L., Harper, G. F., & Mallette, B. (2001). Peer-mediated instruction and interventions and students with mild disabilities. *Remedial and Special Education, 22*, 4–14.
- Maheady, L., Harper, G. F., Mallette, B., & Winstanley, N. (1991). Implementation requirements associated with the use of a classwide peer tutoring program. *Education and Treatment of Children*.
- Maheady, L., Harper, G. F., & Sacca, M. K. (1988). Classwide peer tutoring programs in secondary self-contained programs for the mildly handicapped. *Journal of Research and Development in Education, 21*(3), 76–83.
- Maheady, L., Sacca, M. K., & Harper, G.F. (1988). Classwide peer tutoring with mildly handicapped high school students. *Exceptional Children, 55*, 52–59.
- Mallette, B., Harper, G. F., Maheady, L., & Dempsey, D. (1991). Retention of spelling words acquired using a peer-mediated instructional procedure. *Education and Training in Mental Retardation, 26*(2), 156–164.
- Mallette, B., Maheady, L., & Harper, G. F. (1999). The effects of reciprocal peer coaching on preservice general educators' instruction of students with special learning needs. *Teacher Education and Special Education, 22*, 201–216.
- Shulman, L. S. (2002). Truth and consequences? Inquiry and policy on teacher education. *Journal of Teacher Education, 53*, 248–254.
- Sideridis, G. D., Utley, C. A., Greenwood, C. R., Dawson, H., Delquadri, J. C., & Palmer, P. (1997). An intervention strategy to enhance spelling performance and social interaction and to decrease inappropriate behaviours of students with mild disabilities and typical peers in an inclusive instructional setting. *Research in Education, 59*, 109–124.
- Smith, C. (1987). *MacMillan Spelling*. New York: MacMillan.
- Utley, C. A., Mortweet, S. L., & Greenwood, C. R. (1997). Peer-mediated instruction and interventions. *Focus on Exceptional Children, 29*(5), 1–23.
- Wallace, E., Taylor, W., Fay, L., Kucera, H., & Gonzalez, G. (1988). *Riverside spelling*. Chicago: Riverside Publishing.
- Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education, 53*, 190–204.
- Wright, J. E., 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.

Author information??