

Evaluating the Effectiveness of the CREDE Coaching Model

**A paper presented at the CREDE Conference, San Francisco, CA
May 28-29, 2009**

Presenters:

**Annala Teemant, Indiana University Purdue University Indianapolis
Serena Tyra, University of California, Berkeley
Joan Wink, California State University Stanislaus**

*This research was funded by a grant # P336C050023 from the
U.S. Department of Education Office of Innovation and Improvement*

Evaluating the Effectiveness of the CREDE Coaching Model

Dr. Annela Teemant, Indiana University Purdue University Indianapolis

Dr. Serena Tyra, Educational Consultant

Dr. Joan Wink, California State University Stanislaus

Author Note

Annela Teemant, School of Education, Indiana University Purdue University.

This research was supported by a three-year U.S. Department of Education Teacher Quality Enhancement Grants for State and Partnerships (P336C050023) grant entitled “Improving Teacher Quality from Recruitment to Induction: Integrating Subject Matter, Language and Diversity Pedagogy” and awarded to the Regents of the University of California Berkeley. We would like to acknowledge Hyunchang Kang, a research assistant at the Center for Urban and Multicultural Education, for his assistance in running the statistical analyses.

Correspondence concerning this article should be addressed to Annela Teemant, School of Education, Indiana University Purdue University Indianapolis, 902 West New York Street, ES 3121, Indianapolis, IN 46202. Email: ateemant@iupui.edu

Abstract

This paper describes and evaluates the effectiveness of an instructional coaching model focused on use of five research-based teaching practices promoted by the Center for Research on Education, Diversity and Excellence (CREDE). These pedagogical practices improve academic achievement among diverse student populations. Forty-one elementary teachers participated in a full year of professional development, combining a 30-hour intensive workshop with individual coaching. Using the Standards Performance Continuum, data were gathered during each teacher coaching cycle. Findings demonstrate a significant trend for how teachers incorporate the Five Standards and how instructional coaching significantly improved teacher use of the Five Standards. Improvements in the coaching intervention are needed to support all teachers in reaching the highest levels of implementation.

Evaluating the Effectiveness of the CREDE Coaching Model

The Center for Research on Education, Diversity & Excellence (CREDE) promotes use of five research-based teaching practices, called the Standards for Effective Pedagogy (Five Standards), as essential in improving teacher pedagogy for working with diverse student populations (Tharp, Estrada, Dalton, & Yamauchi, 2000). Multiple studies have demonstrated that when teachers use the Five Standards, they statistically improve student achievement scores on standardized tests (e.g., Doherty & Hilberg, 2007; Doherty, Hilberg, & Lee, 2004; Doherty, Hilberg, Pinal, & Tharp, 2003; Saunders & Goldenberg, 1999).

Under the auspices of a U.S. Department of Education Teacher Quality Enhancement Grant, CREDE (University of California Berkeley), California State University, Stanislaus, and several local school districts in California's Central Valley initiated a partnership for improving how new teachers were pedagogically prepared to teach in high-need schools. One distinctive goal of the grant was to institutionalize use of CREDE's Five Standards at every layer of teacher preparation. The creation of high quality clinical placement sites for new teacher candidates, where use of the Five Standards would be a regular part of best teaching practices, was an essential component of the grant.

This paper describes and evaluates the effectiveness of instructional coaching as a professional development strategy with elementary teachers in creating Five Standards classrooms in these partner districts. Using two years of instructional coaching data, this study explores the nature and quality of change in teacher use of the Five Standards.

Research-Based Practices for Diverse Learners

The Five Standards embody sociocultural principles of learning (Dalton, 1998; Tharp, Estrada, Dalton, & Yamauchi, 2000; Tharp & Gallimore, 1988; Vygotsky, 1978). The Five Standards are a set of observable, research-based teaching practices considered necessary for supporting the academic success of diverse learners. Chart 1 provides a definition for each of the Five Standards.

While each of the Five Standards have individually been shown to have a positive impact on student achievement (e.g., Doherty, Hilberg, & Lee, 2004; Doherty, Hilberg, Pinal, & Tharp, 2003; Estrada, 2004; Estrada & Imhoff, 1999; Saunders & Goldenberg, 1999; Tharp, 1982), the intention is for teachers to use, at least, three of the Five Standards simultaneously in their design of learning activities. In addition, teachers are to employ small group activity centers, rather than whole class instruction, as the primary model of instruction. Together, use of the Five Standards (pedagogy) and small group activity centers (instructional model) make up the Five Standards Instructional Model.

CREDE researchers (Doherty, Hilberg, Epaloose, & Tharp, 2002; Hilberg, Doherty, Epaloose, & Tharp, 2004) developed a five-point rubric, called the Standards Performance Continuum (SPC), to define use of each standard in teaching practice and to highlight the importance of using multiple standards simultaneously in small group activity centers. (The SPC is available at <http://www-gse.berkeley.edu/research/crede/research/tier/spc.html>).

Recent studies have examined Five Standards use and student achievement in naturally occurring classrooms, where teachers have not participated in clearly defined professional

Chart 1

Five Standards for Effective Pedagogy

| |
|--|
| <p>Standard 1: Joint Productive Activity (JPA) Teacher and Students Producing Together <i>Facilitate learning through joint productive activity among teacher and students.</i> Enacting Level: The teacher and a small group of students collaborate on a joint product.</p> |
| <p>Standard 2: Language & Literacy Development (LLD) Developing Language and Literacy Across the Curriculum <i>Develop competence in the language and literacy of instruction across the curriculum.</i> Enacting Level: The teacher provides structured opportunities for students to engaged in sustained reading, writing, or speaking activities; and assists academic language use or literacy development by questioning, rephrasing, or modeling.</p> |
| <p>Standard 3: Contextualization (CXT) Making Meaning: Connecting School to Students' Lives <i>Connect teaching and curriculum to experiences and skills of students' home and community.</i> Enacting Level: The teacher integrates the new activity/information with what students already know from home, school, or community.</p> |
| <p>Standard 4: Challenging Activities (CA) Teaching Complex Thinking <i>Challenge your students toward cognitive complexity.</i> Enacting Level: The teacher designs and enacts challenging activities with clear standards, performance feedback, and assists the development of more complex thinking.</p> |
| <p>Standard 5: Instructional Conversation (IC) Teaching Through Conversation <i>Engage students through dialogue, especially the Instructional Conversation.</i> Enacting Level: The teacher has a planned, goal-directed conversation with a small group of students on an academic topic; elicits student talk by questioning, listening, and responding to assess and assist student understanding; and inquires about students' views, judgments, or rationales. Student talk occurs at higher rates than teacher talk.</p> |

development targeting increased use of the Five Standards (i.e., Doherty & Hilberg, 2007; Doherty, Hilberg, & Lee, 2004; Doherty, Hilberg, Pinal, & Tharp, 2003). Findings demonstrated that native and non-native speakers of English alike made significant gains on standardized tests when teachers used the Five Standards Instructional Model. What has not been investigated is whether a) teachers can be supported in using the Five Standards Instructional Model, and b) how teachers incorporate the Five Standards into their practices.

Description of Instructional Coaching: Performance Against a Standard

CREDE researchers identified instructional coaching as the best professional development strategy to pursue in working with public school teachers because the Five Standards Instructional Model focuses on classroom instruction (Hilberg, Doherty, & Reveles, 2004). The purpose of instructional coaching is to assist teachers in incorporating the Five Standards Instructional Model in their every day teaching practices. Instructional coaching is in harmony with a sociocultural view of teaching and learning. From a sociocultural perspective (Tharp & Gallimore, 1988), coaching provides a “chain of assistance” (p. 83). The coach is the more knowledgeable who assists teacher development by working collaboratively and individually in a teacher’s Zone of Proximal development (ZPD). The coach and teacher use the SPC to guide their coaching interactions. In other words, teacher performance is evaluated against a standard—the SPC.

CREDE’s instructional coaching relies on a two phases of professional development. Phase one consists of participating teachers attending an intensive, five-day, 30-hour workshop prior to the start of job-embedded instructional coaching. The workshop focuses on defining the Five Standards, presenting the use of activity centers, discussing classroom management strategies for phasing in use of activity centers, and designing quality instructional conversations. Phase two consists of extensive individual instructional coaching with each participating teacher across the school year.

The coaching process itself has three stages (Hilberg, Doherty, & Reveles, 2004). For stage one, the coach and teacher meet for 30 minutes to jointly plan a lesson. For stage two, the coach observes the jointly planned lesson for at least 45 minutes, gathering evidence for the follow-up discussion. SPC ratings, teacher and student talk, and notes on the nature of

interactions are scripted throughout the observation at different activity centers. For stage three, the coach and teacher hold a 30-minute debriefing to identify strengths and weaknesses in the lesson. The SPC is used regularly in the planning, observing, and debriefing stages of each coaching cycle to reinforce the performance targets for coaching.

The current study evaluates the effectiveness of instructional coaching as a professional development strategy in supporting teacher use of the Five Standards Instructional Model. This study investigates if and how teachers meet performance targets, as defined by the SPC. Identifying effective professional development models for creating teachers who purposefully meet the learning needs of diverse students through use of research-based practices is a pressing need in schools with diverse student populations. Two main research questions guide this study:

1. Is there an increase in teacher use of the Five Standards as measured by the SPC?
2. What trends exist in how teachers develop in their use of the Five Standards?

Methods

In the context of the Teacher Quality Grant, quantitative data were gathered during the coaching process for the second and third years of the grant. This longitudinal study uses a descriptive pre-experimental design to evaluate the nature and quality of change in elementary teachers' pedagogy using instructional coaching.

Participants

Forty-one elementary teachers (36 females; 5 males) from three high need schools in California's Central Valley participated in a year-long professional development program which combined a 30-hour intensive workshop with a full-year of individual instructional coaching. Eighteen teachers were coached in the second year and 23 teachers were coached in the third

year. The 41 teachers were ethnically diverse: 41.5% White, 43.9% Hispanic, 7.3% Asian, 4.9% Black, and 2.4% identified as other. Thirty-one of the teachers were mainstream teachers, 8 were bilingual teachers, and 2 were specialists (special education and mathematics). Teachers came from every grade level between K-8: K-1 (34%), 2-3 (34%), 4-6 (27%), and mixed grade classes (10%). Teachers were predominately coached during their language arts block (93%).

Three schools, one from each of the three districts, were represented in the database (School 1= 9 teachers; School 2= 26 teachers; and School 3= 6 teachers). Each school has high populations of Hispanic students (Schools 1 & 3= over 84%; School 2= over 70%) and smaller but diverse populations of white, African American, Asian, American Indian, Filipino, and Pacific Islander students. A majority of students are English Language Learners (2007-08: School 1= 55%; School 2= 60% ; School 3= 67%) and students on free or reduced lunch (in 2007-08: School 1= 87%; School 2= 79%; School 3= 54%). School 3, a bilingual school, joined the study in year three of the grant.

The Standards Performance Continuum Observation Tool

The SPC has been described and rigorously tested for reliability and validity (Doherty, Hilberg, Epaloose, & Tharp, 2002; Hilberg, Doherty, Epaloose, & Tharp, 2004; Hilberg, Doherty, Epaloose, & Tharp, 2004). The SPC defines distinct levels of implementation along a continuum, where 0= not observed; 1= emerging (some element present); 2= developing (partial enactment); 3= enacting, meaning the standard is fully enacted as intended; and 4= integrating (at least 3 standards are fully enacted at the same time). The use of each standard can be quantitatively measured. The highest score possible for individual standards is 4. The highest possible total score is 20, with overall levels of enactment defined as follows: 1) emerging < 7.50; 2) developing= 7.50 – 12.49; 3) enacting= 12.50 – 17.49; and 4) integrating= 17.50 – 20.00

(R.S. Hilberg, personal communication, December 12, 2006). Chart 1 (above) provides the description of what is happening for each standard individually when it is fully enacting (i.e., the enacting level) as described on the SPC. To achieve the highest or integrating level on the rubric, a teacher must simultaneously use at least three of the Five Standards at the enacting level in a single activity setting (i.e., the 3 x 3 rule).

Across the two-year study, three instructional coaches were involved in data collection. Two coaches worked with 6 teachers each and the third coach worked with 29 teachers. During a 45-minute observation of teaching performance, the coaches recorded SPC ratings for every individual activity centers. An overall score for the entire observation period is obtained by recording the highest level of Five Standards implementation obtained across centers. In obtaining this overall score, the 3 x 3 rule only applies when enacting scores occur within a single activity setting. For example, it is possible to have an overall score with 3 enacting level ratings for three of the Five Standards not raised to the integrating level rating of 4 if those enacting scores did not occur in the same activity setting.

The Coaching Intervention

Coaching cycles are ordered (e.g., first, second, third) but not necessarily equally spaced across the year. The goal was to engage all teachers in a total of seven coaching cycles. Each cycle consisted of a pre-conference, an observation, and a post-conference. Year-round schedules, district testing, and teacher absences made the number of coaching cycles each teacher received vary, with 40 teachers receiving 4 coaching sessions and only 22 teachers receiving seven coaching cycles.

For coached teachers in this study, the first observation serves as a baseline measure of teacher performance against the standard, which is captured in the SPC rubric. The debrief following the first observation begins the coaching process focused on improving teacher performance. Next steps for teaching were generated in each debrief. The coaching cycles, which were as many as seven, captured teacher development overtime in response to coaching interactions. The final observed lesson documented a teacher's culminating performance against a standard, as defined by the SPC.

Data Analysis

To determine the nature and quality of change in teacher pedagogy elicited through instructional coaching, three steps in data analysis occurred. First, descriptive statistics were run for all independent and dependent variables: namely, frequencies, means, and standard deviations. Second, multiple repeated measures one-way ANOVAs were conducted: Joint productive activity (JPA), language and literacy development (LLD), contextualization (CXT), challenging activities (CA), instructional conversations (IC), and the overall Total Score for all Five Standards. The ANOVA was used to determine if and when changes across coaching cycles (1 to 7) were significantly different. Mauchly's test of sphericity (i.e., tests if variances between means are equal, requiring statistical correction) was not significant for any dependent variables. Reported within-subject effects use the Lower-bound F statistic at $p < .05$. A partial eta squared value was used to identify small ($< .20$), medium ($> .20$ and $< .79$) and large ($> .80$) effect sizes. Third, line graph representations of the development for each standard and the Total Score are provided. Finally, tests of within-subjects contrasts indicated if a significant trend or pattern existed in the data.

Results

The findings for this study address the nature and quality of change in teacher use of the Five Standards Instructional Model elicited through instructional coaching. The findings are reported by research question.

Teacher Use of the Five Standards

Research question one asks whether there is an increase in teacher use of the Five Standards as measured by individual standards (highest possible score= 4) and Total Score (highest possible score= 20) across the coaching process. Table 1 provides the means, standard deviations, and number of subjects for each of the Five Standards and Total Score by coaching cycle. The means for each standard and Total Score consistently increase from one coaching cycle to the next from cycles one to five. During coaching cycle six, the mean teacher performance for each of the Five Standards and Total Score declined slightly. With the exception of CXT, teacher performance improved during coaching cycle seven, returning to a higher mean score than during coaching cycle 5. Standard deviations (SDs) generally increased, with few exceptions, across coaching cycles, showing increasing variation among teachers in their implementation of the standards. Only the SDs for JPA narrowed in cycles six and seven. The SD for IC was the largest of all the Five Standards from coaching cycle one to six; only in coaching cycle 7 did the SD for IC decline, leaving CXT as the standard with the most variation among teachers by the final coaching cycle.

A repeated measures ANOVA revealed there were significant differences in use of the Five Standards, measured by total mean score, between coaching cycle one and seven, $F(1, 21)= 82.85, p < .001$, with a large effect size (eta-squared= .798). For teachers completing seven

Table 1

Means and Standard Deviations for the Five Standards and Total Score by Coaching Cycle

| Five Pedagogical Standards | | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 |
|-------------------------------------|----|---------|---------|---------|---------|---------|---------|---------|
| Joint Productive Activity (JPA) | M | 1.76 | 2.59 | 2.76 | 3.05 | 3.26 | 3.23 | 3.36 |
| | SD | 0.77 | 0.99 | 1.07 | 1.07 | 1.04 | 0.92 | 0.90 |
| | n | 41 | 37 | 41 | 41 | 35 | 31 | 22 |
| Language/Literacy Development (LLD) | M | 2.02 | 2.49 | 2.85 | 3.12 | 3.26 | 3.00 | 3.35 |
| | SD | 0.72 | 0.96 | 0.96 | 0.98 | 0.99 | 1.00 | 0.94 |
| | n | 41 | 37 | 41 | 41 | 34 | 31 | 23 |
| Contextualization (CXT) | M | 1.58 | 2.16 | 2.49 | 2.63 | 3.03 | 2.87 | 2.68 |
| | SD | 0.68 | 1.03 | 1.29 | 1.20 | 1.17 | 1.12 | 1.29 |
| | n | 40 | 38 | 41 | 41 | 34 | 31 | 22 |
| Challenging Activities (CA) | M | 1.66 | 2.19 | 2.54 | 2.78 | 3.03 | 2.97 | 3.18 |
| | SD | 0.62 | 1.00 | 1.05 | 1.06 | 1.06 | 1.11 | 1.05 |
| | n | 41 | 37 | 41 | 41 | 34 | 31 | 22 |
| Instructional Conversation | M | 1.20 | 1.49 | 1.90 | 2.24 | 2.65 | 2.57 | 2.77 |
| | SD | 0.95 | 1.10 | 1.50 | 1.46 | 1.45 | 1.50 | 1.19 |
| | n | 41 | 37 | 41 | 41 | 34 | 30 | 22 |
| Total Score for Five Standards | M | 8.24 | 10.84 | 12.78 | 14.05 | 15.21 | 14.56 | 15.45 |
| | SD | 2.84 | 4.44 | 5.29 | 4.97 | 5.26 | 4.94 | 4.35 |
| | n | 41 | 37 | 41 | 40 | 34 | 32 | 22 |

coaching sessions (n= 22), the mean Total Score at coaching cycle seven (15.45) was significantly greater than the mean Total Score (7.86) for coaching cycle one.

A repeated measures ANOVA revealed there were also significant differences in use of each of the Five Standards individually between coaching cycle one and seven, $F(1, 21)$ JPA= 67.52, LLD= 37.64; CXT= 28.39; CA= 67.52; IC= 40.11. The partial eta squared values indicated large effect sizes for JPA (.76) and CA (.76) and medium effect sizes for LLD (.63), CXT (.58), and IC (.66). For teachers completing seven coaching cycles ($n= 22$), the mean use of each standard at coaching cycle seven (JPA= 3.36; LLD= 3.35 ; CXT= 2.68; CA= 3.18; IC= 2.77) was significantly higher than at coaching cycle one (JPA= 1.77 ; LLD= 2.04 ; CXT= 1.41; CA= 1.59 ; IC=1.05). These findings suggest that the instructional coaching process resulted in significant change in teacher use of the Five Standards, as measured by individual standards and Total Score.

Development Trends

Research question two focuses on what trend or pattern of development emerged across coaching cycles of teacher implementation of the Five Standards Instructional Model. For teachers completing all seven coaching cycles, a repeated measures ANOVA revealed significant differences in Total Score mean across coaching cycles, $F(1, 20)= 19.28, p < .001$, with a medium effect size (partial eta-squared= .49). Table 2 presents the within-subjects contrasts, revealing both a significant and large effect for a linear trend and a significant and medium effect for a single bend or quadratic trend in teacher performance across coaching cycles.

Table 3 presents the Total Score means, standard errors, and 95% confidence intervals for each coaching cycle. The highest Total Score mean (16.62 out of 20 possible) occurred during coaching cycle 5. This is a mean result at the enacting level (12.50 – 17.49) rather than an integrating level (17.50-20.00) of implementation. Figure 1 plots the Total Score means as a line graph. The significant linear trend from cycle one to five is presented with a single quadratic or

Table 2

Within-Subjects Contrasts for Total Score

| Source | factor | Type III Sum of Squares | df | MS | F | Partial Eta-Squared | Power |
|--------------------|-----------|----------------------------|----|--------|--------|------------------------|-------|
| Factor 1 | Linear | 1074.87 | 1 | 633.84 | 74.97* | .79 | 1.00 |
| | Quadratic | 351.12 | 1 | 259.28 | 30.13* | .60 | 0.99 |
| Error (factor1) | Linear | 286.74 | 20 | 7.65 | | | |
| | Quadratic | 233.09 | 20 | 8.77 | | | |

* $p < .001$.

Table 3

Total Score Mean, Standard Error, and Confidence Interval by Coaching Cycle

| Cycle | M | SE | 95% CI | |
|-------|-------|------|--------|-------|
| | | | LL | UL |
| 1 | 8.00 | 0.55 | 6.86 | 9.14 |
| 2 | 10.24 | 0.90 | 8.37 | 12.11 |
| 3 | 13.91 | 1.32 | 11.16 | 16.65 |
| 4 | 16.24 | .87 | 14.42 | 18.06 |
| 5 | 16.62 | 1.09 | 14.35 | 18.89 |
| 6 | 16.10 | 1.00 | 14.01 | 18.18 |
| 7 | 15.81 | .90 | 13.94 | 17.68 |

Note. CI= confidence interval; *LL*= lower limit, *UL*= upper limit.

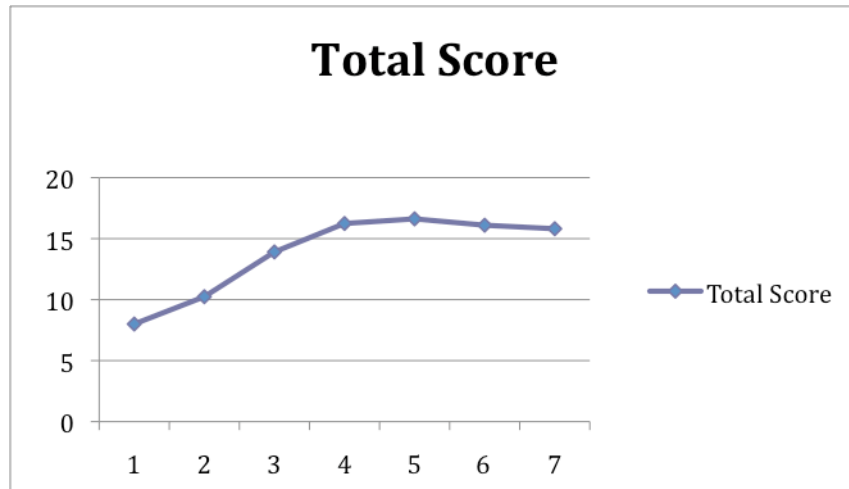


Figure 1. Line representation of mean Total Score cycle to cycle.

downward bend in cycles six and seven, where the increase levels off and falls slightly by the seventh coaching cycle.

Table 4 presents the Total Score mean differences, standard errors, and 95% confidence intervals for statistically significant cycle-to-cycle pairwise comparisons in teacher performance. All Total Score mean differences between cycle one and each subsequent cycle and between cycle two and each subsequent cycle are statistically significant. The only other significant difference in Total Score across coaching cycles occurred in the change between coaching cycle three and five. Mean differences across cycles five, six, and seven did not result in statistically significant improvements in teacher performance based on the analysis of Total Score Mean. Figure 2 presents the mean change cycle to cycle as a bar graph. The greatest amount of change in teacher pedagogy, as measured by Total Score, occurs between cycles one and four. Together these findings demonstrate clear trends in teacher development and change.

Discussion

This study examined the effectiveness of instructional coaching that relied on a set of performance-based targets captured by the Five Standards Instructional Model. The study tested

the hypothesis that instructional coaching would lead to increased use of the Five Standards Instructional Model by teachers. This study also sought to describe trends in teacher development in relationship to implementation of the Five Standards.

Table 4
Mean Differences, Standard Errors, and Confidence Intervals for
Total Score Pairwise Comparisons

| Cycles | | Mean Difference (A-B) | SE | 95% CI | |
|--------|---|--------------------------|------|--------|-------|
| A | B | | | LL | UL |
| 1 | 2 | -2.24* | 0.81 | -3.94 | -.54 |
| | 3 | -5.91* | 1.31 | -8.64 | -3.17 |
| | 4 | -8.24* | 0.95 | -10.22 | -6.25 |
| | 5 | -8.62* | 1.00 | -10.70 | -6.54 |
| | 6 | -8.10* | 0.94 | -10.06 | -6.13 |
| | 7 | -7.81* | 0.84 | -9.57 | -6.05 |
| 2 | 3 | -3.67* | 0.99 | -5.73 | -1.60 |
| | 4 | -6.00* | 1.07 | -8.23 | -3.77 |
| | 5 | -6.38* | 1.17 | -8.82 | -3.95 |
| | 6 | -5.86* | 1.10 | -8.14 | -3.57 |
| | 7 | -5.57* | 1.10 | -7.87 | -3.27 |
| 3 | 5 | -2.71* | 1.25 | -5.32 | -.011 |

Note. CI= confidence interval [lower limit, upper limit]. * $p < .05$

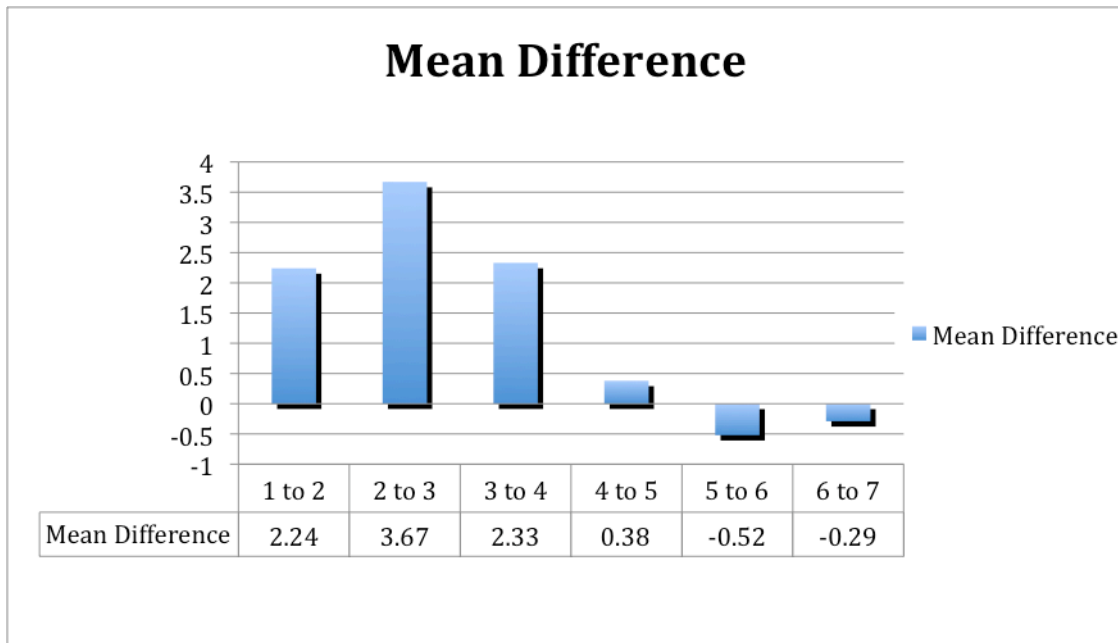


Figure 2. Bar graph representation of mean change in Total Score cycle to cycle.

Findings indicate that Five Standards instructional coaching, where the SPC provides the standards against which teacher performance is measured, is an effective professional development strategy for working with teachers in public school settings. Instructional coaching leads to statistically significant increases in teacher use of the Five Standards measured individually and as Total Score from coaching session one to seven. Coaching also leads to a statistically significant overall trend of greater use of the Five Standards over coaching cycles, with peak performance occurring at coaching cycle five.

Two theoretical implications emerge from the findings. First, the coaching outcomes indicate that the instructional coaching process itself is effective in assisting teachers in transforming their pedagogy to meet the learning needs of diverse learners. Each coaching conversation resulted in teacher improvement, with the most rapid, linear, and significant change in teacher pedagogy occurring during the first five coaching cycles. In the context of this study,

the coach-teacher dialogue begins with co-planning, moves to observable teaching actions, and then ends with reflection upon teaching actions in light of defined targets. Such instructional coaching, unlike other professional development strategies, embodies both Vygotsky's (1978) sociocultural perspectives and Friere's (1994) notion of praxis (e.g., learning through engaging in action and reflection upon the outcomes of action). Instructional coaching that values learning as a social- and action-oriented process promotes measureable and meaningful transformation in teacher pedagogy.

Second, the coaching outcomes also demonstrate that it is useful to coach with defined targets for teacher performance in mind. The SPC, in particular, provides concrete targets against which both teacher performance and coaching outcomes can be evaluated. The SPC responds to the question teachers and coaches often ask, "Where should I begin to improve teaching?" The Five Standards Instructional Model, as the content of the coaching process, results in goal-directed and performance-based transformation. The findings provide quantitative evidence of which targets for teacher performance are incorporate into teaching practice earlier or later in the coaching process. For example, teachers are most inclined to incorporate into their teaching (a) an increased focus on collaborative products (JPA), (b) greater language use by students (LLD), and (c) clearer expectations with higher cognitive challenge guiding student work (CA). Teachers move more slowly to incorporate rich contextualization (CXT) and the instructional conversation (IC).

Three practical concerns and implications emerge from these findings for improving the coaching process. First, although all coached teachers experienced significant growth in use of the Five Standards Instructional Model, teachers did not collectively reach in the highest level of implementation. In other words, while some teachers were able to use three of the Five Standards

simultaneously, (enacting level= 12.50 <17.49), a majority of teachers did not reach the integrating level (17.50 -20.00) of the SPC, where three or more of the Five Standards would be used in the same activity setting. While there is general fidelity to the instructional model, there is room for improvement in promoting higher overall fidelity.

Second, the standards of CXT and the IC are the two most challenging standards for teachers to implement and sustain at the highest level. For example, teachers are able to make incidental connections between students' lives and new academic content, but they fail to fully integrate students' previous knowledge into the learning of new academic content. Although teachers use small group activity centers and are able to transition to working with students at a single activity center, they still struggle to meet the requirements of the instructional conversation. Student talk rather than teacher talk should dominate and student rationales for thinking should be elicited as a part of the instructional conversation. Findings suggest that CXT and the IC present unique challenges to coaches and teachers alike. The final coaching cycles, in particular, should be used differently and/or more effectively to assist teachers to implement these more challenging standards. Improving teacher use of CXT and IC would also improve overall fidelity to the model.

Third, the quadratic trend in teacher development indicates a decline in teacher growth by the sixth and seventh coaching cycles. The sixth coaching cycle occurred at the end of February, just prior to the beginning of standardized testing in early March both years of the grant. The decline in teacher growth could reflect the timing of this coaching cycle and/or pressure teachers feel to teach in ways perceived as preparing students to take standardized tests. For coaching cycle seven, half of the teachers were coached during March and the other half during May. Because there is a slight improvement in teacher scores between the sixth and seventh coaching

cycle, it remains unclear, whether the decline in teacher growth in these final coaching cycles is the result of timing, the coaching process, or unidentified issues in teacher development related to implementing the standards. Nevertheless, one clear implication for coaching is to work around standardized testing schedules more directly to investigate if this leads to improved outcomes for the final coaching cycles as well as overall fidelity to the model.

Further qualitative and quantitative research is needed to understand the challenges associated with implementing use of the Five Standards Instructional Model. For example, what are the perceived difficulties or barriers for teachers in contextualizing lessons or using the instructional conversation more effectively? Are there student, teacher, curriculum, or policies factors that limit the quality and quantity of teacher change related to specific standards? In addition, further research is needed to investigate coached achievement in contrast to teacher sustainability. While the highest level of implementation occurred at coaching cycle five in this study, the findings do not indicate if or how many coaching cycles are needed to ensure sustained use of the Five Standards Instructional Model over time. Multi-year studies, which follow teacher performance a year after coaching, would allow deeper understanding of the obstacles and/or the required conditions for accomplishing sustained teacher use of the Five Standards Instructional Model.

Conclusion

Preparing and retaining teachers to work more effectively with students from increasingly diverse cultural, linguistic, and economic backgrounds are pressing concerns in high-need schools. This paper described a specific performance-based instructional coaching model for assisting teachers to use five research-based practices known to increase student achievement.

The research results provide evidence that performance-based instructional coaching is an effective professional development strategy in promoting teacher change. In addition, the findings demonstrate that teachers can be coached to use the Five Standards Instructional Model with fidelity. When university-public school partners collaborate in the preparation of high quality clinical placement sites, there are reciprocal benefits in the teacher education enterprise for modeling, reinforcing, and demonstrating the effectiveness of research-based practices for new teacher candidates. The use of an instructional coaching process and specific performance targets for teacher growth, such as the Five Standards, hold promise for preparing teachers to work effectively in high need schools.

Reference

- Dalton, S. S. (1998). *Pedagogy matters: Standards for effective teaching practice*. Santa Cruz, CA: University of California, Center for Research on Education, Diversity & Excellence.
- Doherty, R. W., & Hilberg, R. S. (2007). Standards for effective pedagogy, classroom organization, English proficiency, and student achievement. *Journal of Educational Research, 101*(1), 24-35.
- Doherty, R. W., Hilberg, R. S., Epaloose, G., & Tharp, R. G. (2002). Standards performance continuum: Development and validation of a measure of effective pedagogy. *Journal of Educational Research, 96*(2), 78-89.
- Doherty, R. W., Hilberg, R. S., & Lee, V. (2004). *Observing the five standards in practice: Development and application of the standards performance continuum* (CREDE Research Brief #11). Retrieved from Center of Applied Linguistics website: www.cal.org/crede/pdfs/rb11.pdf

- Doherty, R. W., Hilberg, R. S., Pinal, A., & Tharp, R. G. (2003). Five standards and student achievement. *NABE Journal of Research and Practice*, 1(1), 1-24.
- Estrada, P. (2004). Patterns of language arts instruction activity: Excellence in first and fourth grade culturally and linguistically diverse classrooms. In H. C. Waxman, R. G. Tharp, & R. S. Hilberg (Eds.), *Observational research in U.S. Classrooms: New approaches for understanding cultural and linguistic diversity*. Cambridge, MA: Cambridge University Press.
- Estrada, P., & Imhoff, B. (1999). *Patterns of instructional activity: Excellence, inclusion, fairness, and harmony in six first grade classrooms* (Technical Report No. 3). Santa Cruz, CA: University of California, Center for Research on Education, Diversity & Excellence (CREDE).
- Freire, P. (1994). *Pedagogy of the oppressed*. New York: Continuum.
- Hilberg, S., Doherty, W., Epaloose, G., & Tharp, R.G. (2004). The standards performance continuum: A performance-based measure of the Standards for Effective Pedagogy. In H. C. Waxman, R. G. Tharp, & R. S. Hilberg (Eds.), *Observational research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity*, (pp. 49-71). New York, NY: Cambridge University Press.
- Hilberg, S., Doherty, W., & Reveles, C. (2004). *Training manual and resources for five standards coaching*. Unpublished manuscript, University of California, Santa Cruz.
- Saunders, W. M., & Goldenberg, C. (1999). *The effects of instructional conversations and literature logs on the story comprehension and thematic understanding of English*

proficiency and limited English proficient students. Santa Cruz, CA: Center for Research on Education, Diversity & Excellence.

Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha Early Education Program. *Reading Research Quarterly*, 17(4), 503-527.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York, NY: Cambridge University Press.

Tharp, R. G., Estrada, P., Dalton, S. S., & Yamauchi, L. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder: Westview Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*, (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Trans. & Ed.). Cambridge, MA: Harvard University Press.