# Effects of coaching on teacher use of sociocultural instructional practices

Annela Teemant [a,*], Joan Wink [b], Serena Tyra [c]

[a] School of Education, Indiana University Purdue University Indianapolis, 902 West New York Street, ES 3121, Indianapolis, IN 46202, USA
[b] California State University Stanislaus, Department of Teacher Education, DBH 319, One University Circle, Turlock, CA 95382, USA
[c] 1022 Stanford Dr., Simi Valley, CA 93065, USA

A B S T R A C T

This study evaluates a performance-based instructional coaching model intended to improve teacher pedagogy and classroom organization for educating diverse student populations. Elementary teachers ($N = 21$) participated in a 30-h workshop and seven individual coaching sessions across an academic year. The coaching model promoted use of the Standards for Effective Pedagogy, five research-based practices known to increase student achievement. Findings demonstrate performance-based instructional coaching led to statistically significant (a) improvements in teacher pedagogy, (b) patterns of teacher growth, and (c) changes in classroom organization. Implications for improving teachers' ultimate achievement, the coaching protocol, and research are addressed.

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

In the United States, and countries such as Australia, Canada, and England, today's teachers are serving an increasingly diverse student population (e.g., Bernhard, Lefebvre, Chud & Lange, 1997; Cobbold, 2007, 2010; Skerrett, 2008; The Sutton Trust, 2010). In the U.S., 45% of students are children of color, 21% speak a language other than English at home, and 17% of all public schools are high-poverty schools (NCES, 2010). Increasingly, teachers in many countries are being challenged to educate with equity students who are racially, culturally, linguistically, economically, or geographically diverse (e.g., Figueroa, 2004; Hill & Allan, 2004; Hollins & Guzman, 2005; Haworth, 2009).

Unfortunately, many classroom teachers do not feel adequately prepared to teach students of such diversity, as evidenced by teacher flight from the profession (e.g., Barmby, 2006; Ingersoll, 2004; Mackel, 2002). Professional development is a "big ticket item" in school and state operating budgets (Zepeda, 2008, p. 31). Demonstrating efficacy of a professional development model specifically targeting improved teacher performance with diverse learners is a pressing and shared international need (e.g., Cobbold, 2007; Darling-Hammond & Sykes, 1999; Wang, Coleman, Coley, & Phelps, 2003). However, very little research exists documenting the effectiveness of such professional development (Borko, 2004; Knight & Wiseman, 2005).

The Center for Research on Education, Diversity and Excellence (CREDE) has articulated five sociocultural principles essential for teachers working with diverse learners (Tharp, Estrada, Dalton, & Yamauchi, 2000). These principles, as defined in Fig. 1, comprise the Standards for Effective Pedagogy (Five Standards): Joint Productive Activity (JPA), Language and Literacy Development (LLD), Contextualization (CTX), Cognitive Challenge (CA), and Instructional Conversation (IC). Numerous experimental, quasi-experimental, and longitudinal studies have validated the positive effect on student achievement of these standards individually and as an instructional model (e.g., Doherty & Hilberg, 2007; Doherty, Hilberg, Pinal, & Tharp, 2003; Estrada, 2005; Estrada & Imhoff, 1999; Hilberg, Tharp, & DeGeest, 2000; Saunders & Goldenberg, 1999; Tharp, 1982).

Recent studies have used a quantitative classroom observation tool called the Standards Performance Continuum to measure the degree to which teachers are employing the Five Standards in their teaching (Doherty, Hilberg, Epaloose, & Tharp, 2002; Hilberg, Doherty, Epaloose, & Tharp, 2004). In particular, studies by Doherty et al. (2003) and Doherty and Hilberg (2007) provide evidence that teachers, who naturally use the Five Standards and small group activity centers as a classroom organization,

* Corresponding author. Tel.: +1 317 274 1228; fax: +1 317 274 6864.
E-mail addresses: ateemant@iupui.edu (A. Teemant), jwink@csustan.edu (J. Wink), sktyra@aol.com (S. Tyra).

# Standards for Effective Pedagogy

## Standard 1: Joint Productive Activity (JPA)

**Teacher and Students Producing Together**
*Facilitate learning through joint productive activity among teacher and students.*
Enacting Level: The teacher and a small group of students collaborate on a shared product.

## Standard 2: Language and Literacy Development (LLD)

**Developing Language and Literacy Across the Curriculum**
*Develop competence in the language and literacy of instruction across the curriculum.*
Enacting Level: The teacher provides structured opportunities for students to engage in sustained reading, writing, or speaking activities; and assists academic language use or literacy development by questioning, rephrasing, or modeling.

## Standard 3: Contextualization (CTX)

**Making Meaning: Connecting School to Students' Lives**
*Connect teaching and curriculum to experiences and skills of students' home and community.*
Enacting Level: The teacher integrates the new activity/information with what students already know from home, school, or community.

## Standard 4: Challenging Activities (CA)

**Teaching Complex Thinking**
*Challenge students toward cognitive complexity.*
Enacting Level: The teacher designs and enacts challenging activities with clear standards and performance feedback, and assists the development of more complex thinking.

## Standard 5: Instructional Conversation (IC)

**Teaching Through Conversation**
*Engage students through dialogue, especially the Instructional Conversation.*
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.

**CREDE** | Center for Research on Education, Diversity & Excellence (CREDE) | http://crede.berkeley.edu
Graduate School of Education | University of California | Berkeley, CA 94720

**Fig. 1.** The standards for effective pedagogy.

statistically improve student achievement and English proficiency among both native and non-native speakers of English. What is lacking, to date, is research on an effective professional development strategy promoting teacher use of these research-based practices.

The purpose of this paper is two fold. First, it describes a new instructional coaching model that is sociocultural in its process and performance-based in its use of the Five Standards Instructional Model. Second, this paper describes patterns of change in teacher use the Five Standards Instructional Model across time as a result of

instructional coaching. This study contributes quantitative and longitudinal evidence to the growing body of research on instructional coaching as a value-added professional development strategy.

## 2. Relevant literature

Literature relevant to coaching, instructional coaching, and sociocultural theory and pedagogy situate the current study. The research questions guiding this study conclude the section.

### 2.1. Coaching research

Improving teacher quality is pivotal to improving student achievement (e.g., Darling-Hammond, 2000). Coaching has emerged as an effective strategy for ongoing teacher development (e.g., Cornett & Knight, 2009; Joyce & Showers, 1995; Knight, 2009a; Sparks & Hirsh, 1997; Speck & Knipe, 2001).

Coaching is job-embedded, teacher directed, school based, collaborative, ongoing, information rich professional development that focuses on student learning (e.g., Garet, Porter, Desimone, Birman, & Yoon, 2001; Hawley & Valli, 1999; Zepeda, 2008). According to Knight (2009b, p. 1), coaching is being advocated "on a grand scale" by districts and states as "the most promising" strategy for improving schools and student achievement. Coaching puts teachers' needs "at the heart of professional learning by individualizing," (Knight, 2009b, p. 2). Current Race to the Top policy specifically identifies "providing relevant coaching" as a strategy for improving teacher effectiveness (U.S. Department of Education, 2009, p. 9).

In their review of coaching research, Speck and Knipe (2001, p. 88) argue that, "Using teacher coaching as a follow up to new learning has increased the implementation of new concepts and strategies by teachers" and "improved student achievement." In Cornett's and Knight's (2009) narrative review of four prominent educational coaching models (i.e., peer, cognitive, literacy, and instructional coaching), they argue coaching research demonstrates (a) a positive impact on teacher attitudes, (b) increased implementation or skill transfer; (c) increased feelings of teacher efficacy; and (d) improved student achievement.

Cornett and Knight (2009) argue for more empirical research on coaching efficacy. To date, coaching research has not established which coaching practices, with what frequency, promoting what content and teaching practices are the most effective given limited human and financial school resources. Evaluating the effectiveness of professional development is essential in guiding school reform in an era demanding greater accountability, especially when accountability is focused so squarely on teacher performance (Guskey, 2000; U.S. Department of Education, 2009).

### 2.2. Instructional coaching

Instructional coaching, as a special type of educational coaching, takes place in teacher classrooms and targets teacher performance during instruction (Knight, 2004; Kowal & Steiner, 2007). An instructional coach can be a peer, a veteran teacher, or a consultant external to the system.

Knight (2009c) defines instructional coaching as a partnership between a coach and teacher, where there are commitments to (a) equality in the relationship, (b) teacher choice in the content and process of learning, (c) empowerment and respect for varying perspectives, (d) authentic dialogue, (e) reflection, (f) praxis (i.e., reflection and action), and (g) reciprocity of learning between the coach and teacher. Whereas mentors share expertise in one direction, sharing in coaching is "characterized by parity and bidirectionality," (Sherris, 2010, p. 1). In order to accomplish this type of partnership, a coach holds small group meetings for consensus building, unpacks and models specific and desired teaching practices, and then individually interviews teachers, observes teacher implementation, and engages in reflective conversations intended to analyze teacher performance.

Knight (2009c) states that typically the focus of instructional coaching includes classroom management, academic content, instruction, and/or formative assessment. With certainty, the desired outcomes of instructional coaching are site negotiated and partnership specific, based in "challenges present in the classroom," (Zepeda, 2008, p. 170).

The coach–teacher dialogue must in fact facilitate meaningful improvements in teaching. One of the challenges of any instructional coaching model is to identify, define, and then evaluate what constitutes meaningful improvement. While starting points for coaching conversations are easy to identify (e.g., classroom management, content knowledge, etc.), it is much more difficult to provide valid and reliable evidence of improvement over time. For example, research on Knight's model of instructional coaching has relied on self-report data and no "externally obtained reliable data to show whether teachers were actually doing what they report" (Cornett & Knight, 2009, p. 208). More rigorous accountability is required to establish instructional coaching as a value-added professional development strategy.

### 2.3. Why a sociocultural perspective on instructional practices

Sociocultural theory is widely recognized as foundational in teacher education for diversity (Lewis, Enciso, & Moje, 2007; Rogoff, 1990; Rogoff & Wertsch, 1984; Tharp & Gallimore, 1988; Vygotsky, 1978). When teachers and students do not share common cultural, historical, political, or community experiences, it is essential to create a shared context for learning (Cochran-Smith & Zeichner, 2005; Sleeter, 2008).

Vygotsky (1978) framed learning as a social process: It is through dialogue, that teachers—known as more knowledgeable others—assess and assist student learning within the student's Zone of Proximal Development (ZPD). The ZPD embodies the difference between what a student can do on his/her own and what can be done with assistance from a more knowledgeable other (e.g., peer, parent, teacher, etc.). Tharp and Gallimore (1988) argue that teaching provides a "chain of assistance" that advances student learning. For Vygotsky (1997), learning is an active process for the student, the teacher, and in the environment between the student and teacher.

The Five Standards Instructional Model (Dalton, 2008; Tharp et al., 2000) represents a sociocultural pedagogy designed to maximize learning for diverse students. This model is comprised of two parts: principles of learning (i.e., the Five Standards) and a small group configuration of the classroom (i.e., instructional model). The Five Standards, as defined in Fig. 1, call for learning to be a language and literacy rich conversation between teachers and students. That process allows teachers (a) to listen to student thinking, (b) to assess, assist, and influence development through questioning, rephrasing, modeling, and cognitively challenging students, while (c) deeply connecting new learning to students' knowledge from home, school, or community. When taken together, the Five Standards operationalize what it means for a teacher to work in a student's ZPD during instruction (J. Mahiri, personal communication, September 21, 2007).

This instructional model is defined by the existence of both a teacher center—where an instructional conversation takes place between a teacher and students with homogenous learning needs (e.g., Saunders & Goldenberg, 1999)—and use of multiple and

diversified independent activity centers, where heterogeneously grouped students work collaboratively to learn (Tharp et al., 2000). Small group instruction dramatically increases the quality and quantity of opportunities for students to receive teacher and peer assistance in the process of learning (Tharp & Gallimore, 1988; Vygotsky, 1978). Tharp et al. (2000) argue that the goal of achieving excellence, fairness, inclusion, and equity in the classroom hinges on concrete teacher actions embodied in the Five Standards Instructional Model.

For professional development and research purposes, CREDE researchers developed the five-point Standards Performance Continuum (SPC) observation rubric. The SPC measures quantitatively the quality of teacher implementation of the Five Standards Instructional Model (Doherty et al., 2002; Hilberg, Doherty, Epaloose et al., 2004). The continuum is anchored on one end by practices that hold little hope for engaging diverse students: typically whole class, teacher-dominated instruction where individualistic, decontextualized, and rote conceptions of learning dominate. At the opposite end, social interaction and negotiation, assistance and feedback, rich contextualization, collaboration, cognitive challenge, and dialogue are valued. Fig. 2 contains the SPC rubric, which defines concrete teacher actions that move a teacher toward intentional use of sociocultural practices.

In summary, studies have shown that teachers who use the Five Standards Instructional Model possess the instructional and cultural knowledge needed to participate competently in advancing learning among diverse students. The model, therefore, has demonstrated value in serving as research-based performance targets for teaching. Further investigation is needed to determine if, and to what degree, teachers can improve in implementing such practices through defined, targeted, and systematic professional development

## 2.4. Defining process and targets for instructional coaching

The purpose of this study is to define and evaluate a sociocultural professional development model of instructional coaching that supports teachers in meeting the needs of their diverse learners. We advance a sociocultural perspective on instructional coaching, where research-based principles (i.e., Five Standards) define both the process and performance targets for coaching.

### 2.4.1. The sociocultural process

When sociocultural tenets inform the instructional coaching process, then effective instructional coaching occurs when (a) learning is social; (b) coaching is assisting; (c) teacher performance—and therefore coaching— leads development, being dependent on individual needs and specific teaching situations; and (d) knowledge is defined as more cultural understanding and competent participation in student learning.

When coaching focuses on instructional practices, the coach serves as a more knowledgeable other, who assists and collaboratively develops a lesson plan with a teacher, provides data-rich feedback following an observation, and then engages in cycles of

| | NOT OBSERVED | EMERGING | DEVELOPING | ENACTING | INTEGRATING |
|---|---|---|---|---|---|
| *General Definition:* | *The standard is not observed.* | *One or more elements of the standard are enacted.* | *The teacher designs and enacts activities that demonstrate a partial enactment of the standard.* | *The teacher designs, enacts, and assists in activities that demonstrate a complete enactment of the standard.* | *The teacher designs, enacts, and assists in activities that demonstrate skillful integration of multiple standards simultaneously.* |
| Joint Productive Activity<br><br>*Teacher and Students Producing Together* | Students work independently of one another. | Students are seated with a partner or group, AND (a) collaborate or assist one another, OR (b) are instructed in how to work in groups, OR (c) contribute individual work, not requiring collaboration, to a joint product. | The teacher and students collaborate on a joint product in a whole-class setting, OR students collaborate on a joint product in pairs or small groups. | The teacher and a small group of students collaborate on a joint product. | The teacher designs, enacts, and collaborates in joint productive activities that demonstrate skillful integration* of multiple standards simultaneously. |
| Language & Literacy Development<br><br>*Developing Language and Literacy Across the Curriculum* | Instruction is dominated by teacher talk. | (a) The teacher explicitly models appropriate language; OR (b) students engage in brief, repetitive, or drill-like reading, writing, or speaking activities; OR (c) students engage in social talk while working. | The teacher provides structured opportunities for academic language development in sustained reading, writing or speaking activities. (Sustained means at least 10 minutes. If it is a whole class arrangement, then more than 50% of the students are participating. No turn taking.) | The teacher designs and enacts instructional activities that *generate* language expression and development of 'content vocabulary,' AND *assists student language use or literacy development* through questioning, rephrasing, or modeling. | The teacher designs, enacts, and assists in language development activities that demonstrate skillful integration of multiple standards simultaneously. |
| Contextualization<br><br>*Making Meaning – Connecting School to Students' Lives* | New information is presented in an abstract, disconnected manner. | The teacher (a) includes some aspect of students' everyday experience in instruction, OR (b) connects classroom activities by theme or builds on the current unit of instruction, OR (c) includes parents or community members in activities or instruction, OR (d) connects student comments to content concepts. | The teacher makes incidental connections between students' prior experience/knowledge from home, school, or community and the new activity/academic concepts. | The teacher integrates the new activity/academic concepts with students' prior knowledge from home, school, or community to connect everyday and schooled concepts. | The teacher designs, enacts, and assists in contextualized activities that demonstrate skillful integration of multiple standards simultaneously. |
| Challenging Activities<br><br>*Teaching Complex Thinking* | Activities rely on repetition, recall, or duplication to produce factual or procedural information. | The teacher (a) accommodates students' varied ability levels, OR (b) sets and presents quality standards* for student performance, OR (c) provides students with feedback on their performance. | The teacher designs and enacts 'challenging activities' that connect instructional elements to academic content OR advance student understanding to more complex levels. | The teacher designs and enacts challenging activities with clear standards/expectations and performance feedback, AND assists* the development of more complex thinking. | The teacher designs, enacts, and assists in challenging activities that demonstrate skillful integration of multiple standards simultaneously. |
| Instructional Conversation<br><br>*Teaching Through Conversation* | Lecture or whole-class instruction predominates. | With individuals or small groups of students, the teacher (a) responds in ways that are comfortable for students, OR (b) uses questioning, listening or rephrasing *to elicit student talk*, OR (c) converses on a nonacademic topic. | The teacher converses with a small group of students on an academic topic AND *elicits student talk* with questioning, listening, rephrasing, or modeling. | The teacher: designs and enacts an instructional conversation (IC) with a clear academic goal; listens carefully to assess and assist student understanding; AND questions students on their views, judgments, or rationales. Student talk occurs at higher rates than teacher talk. | The teacher designs, enacts, and assists in instructional conversations that demonstrate skillful integration of multiple standards simultaneously. |

**Fig. 2.** Standards performance continuum: a classroom observation rubric.

reflection and action to support implementation of new practices. Coaching, therefore, is a series of instructional conversations between a coach and teacher (Hilberg, Doherty, & Reveles, 2004), where knowledge of practice is co-constructed through shared expertise (Wink & Putney, 2002). Each coaching interaction is goal-directed, dialogic, collaborative, and reflective, allowing the teacher to receive timely, individualized, and meaningful assistance for improving teaching. The use of pre- and post-observation instructional conversations between the coach and teacher allow theoretical and practical knowledge to inform each other in a personally and socially meaningful context: The teacher's classroom.

Instructional coaching, defined through a sociocultural lens, focuses more attention on working within a teacher's zone of proximal development. With access to a more knowledgeable coach, teachers engage in cycles of reflection and action—praxis (Freire, 1994)—that assist and ideally quicken professional growth beyond what the teacher could accomplish alone. Instructional coaching supports the notion that the "ZPD is not a place at all; it is an activity…expressed as revolutionary activity" (Wink & Putney, 2002, p. 153).

### 2.4.2. The sociocultural performance targets

While studies validate the Five Standards Instructional Model (e.g., Doherty & Hilberg, 2007; Doherty et al., 2002), studies to date have not evaluated an instructional coaching model intended for broad use by school districts. In this study, the Standards Performance Continuum rubric defines performance targets for teaching and evaluation of teacher progress. In essence, the rubric ratings reveal a teacher's ZPD in relationship to the model and become the site of expert–novice activity (i.e., conversation, planning, teaching, and reflecting) for the coach and teacher. The evidence gathered by the coach makes teaching improvement visible and actionable. The reflection makes it personally relevant to the teacher.

Crane (2002) argues that for coaching to be transformative, it must carefully balance measurable outcomes—desired performance—with quality relationships. The coach–teacher partnership revolves around making progress toward increased use of sociocultural practices; however, the rate, route, and ultimate attainment of these performance targets remain teacher directed. Authentic transformation must be developmentally appropriate and intrinsically satisfying to the teacher, not forced. Therefore, while the SPC provides the performance targets for teaching, the reflective conversations are individualized. Conversations may focus on management, academic content, instructional strategies, analysis of student work, or a combination of these topics, but they are simultaneously anchored by the Five Standards Instructional Model as an agreed upon performance target.

### 2.5. Research questions

This study examines the value of performance-based instructional coaching when the Five Standards Instructional Model defines performance. Teachers' fidelity to the model is expected to increase based on the use of instructional coaching. The number of coaching cycles is the independent variable, and the dependent variables include each of the Five Standards, the total SPC score, number of activity centers, phase of classroom organization, and quality of the instructional conversation. The research questions are:

RQ1. Is there an increase in teacher use of the Five Standards as measured by individual standards and total score?

RQ2. What patterns of implementation for the Five Standards Instructional Model emerge across coaching cycles for all teachers, and do patterns of implementation significantly vary based on teachers' starting points for learning (i.e., initially high or low group implementers)?

RQ3. Is there a change in classroom organization as indicated by phase of classroom organization and number of small group activity centers?

## 3. Methods

This study is descriptive, quantitative, and longitudinal. It uses a repeated measures design to evaluate (a) if, (b) to what degree of fidelity, and (c) with what pattern coached teachers enacted the Five Standards Instructional Model. A description of the participants, instruments, and intervention follows.

### 3.1. Participants

Participants are drawn from a larger database (Teemant, Tyra, & Wink, 2009). The selection criterion for inclusion in this study was completion of seven coaching cycles. Twenty-one teachers (2 male) from two elementary schools (School 1 = 16 teachers; School 2 = 5 teachers) were selected. Teachers were experienced (3–20 years), ethnically diverse (43% White, 33% Hispanic, 14% Asian, and 10% Black), and represented each grade level: K-1 (38%), 2–3 (33%), 4–6 (19%), and one mixed grade 4/5 classes (5%).

Each school has high populations of Hispanic students (School 1 over 70% and School 2 over 84%) and smaller but diverse populations of White, African American, Asian, American Indian, Filipino, and Pacific Islander students. The schools also have high numbers of English Language Learners (2007–2008: School 1 = 60%; School 2 = 67%) and students on free or reduced lunch (in 2007–2008: School 1 = 79%; School 2 = 54%). School 2 is a bilingual school.

### 3.2. The standards performance continuum

The SPC is a valid and reliable observation rubric (Doherty et al., 2002; Hilberg, Doherty, Epaloose et al., 2004). It quantitatively measures teacher use of the Five Standards. Each standard is measured along a 5-point 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, which can only be achieved when no less than three of the five (3 × 3 rule) standards are fully enacted at the same time in a single activity. The highest total score possible is 20. Small group configurations result in higher SPC scores. Hilberg (Personal communication, December 12, 2006) provided four value ranges for determining fidelity of implementation: 1) emerging < 7.50; 2) developing = 7.50–12.49; 3) enacting = 12.50–17.49; and 4) integrating = 17.50–20.00.

Three instructional coaches served as external consultants with expertise in the Five Standards Instructional Model (Coach 1 = 6 teachers; Coach 2 = 5 teachers; Coach 3 = 11). Their SPC ratings illuminated the cycle-to-cycle pattern of teachers' development. They received training in use of the SPC by its developers in 2002–2003. Two coaches participated in a five-day, inter-rater reliability seminar with ten participants rating video clips (Kendall's *W*: Joint Productivity = .96; Language/Literacy = .98; Contextualization = .83; Challenging Activities = .88; Instructional Conversation = .95). The other coach received training one-on-one while coaching. In 2005, the three coaches spent four months calibrating their use of the SPC in coaching to achieve consensus on ratings. While no inter-rater reliability data for all three coaches was gathered in 2005, an ANOVA revealed no rater bias in a model with Total Score, Coach, and initially High (<8)/Low (≤8) Group Implementers: $F(2, 21) = 2.37$, $p = .13$.

### 3.3. The Instructional Coaching Procedures, Intervention, and Targets

The professional development relied on two phases of activity conducted by the external coaches. In phase one, teachers attended an intensive, five-day, 30-h workshop focused on defining the Five Standards, the activity center classroom organization, a 8–12 week phase-in process, and quality instructional conversations (Hilberg, Chang, & Epaloose, 2003). Teachers learned how to systematically teach norms, expectations, and procedures for successful group collaboration, preparing students to work independent of the teacher. At any given time, curricular content and instructional strategies influence classroom organization; however, instructional coaching in this study purposefully targeted developing teacher expertise in using multiple, simultaneous, and diversified activity centers successfully.

Phase two consisted of individual instructional coaching during language arts across a school year (August to May; approximately 15 contact hours). The seven coaching cycles were ordered (e.g., first, second) rather than equally spaced events. Coaches used observation records to capture: (a) overall quality of Five Standards use; (b) number and quality of individual activity centers; and (c) whether the teacher relied primarily on whole class organization, student-led small group activity centers, or the addition of a teacher center with an instructional conversation as one of several diversified activity centers.

Prior to coaching, the coach and teacher met for an interview to build rapport and create shared goals and expectations. This first meeting, immediately followed by a baseline observation, allowed the coach to establish where the teacher was developmentally on the SPC. The final observed lesson documents a teacher's culminating SPC performance.

The coaching process itself has three stages (Hilberg, Doherty, & Reveles, 2004). For stage one, the coach and teacher meet for 30 min to review a planned lesson together. The discussion is teacher driven and coach facilitated, with focus on phasing in use of the Five Standards Instructional Model. For stage two, the coach observes the planned lesson for at least 45 min, gathering evidence for the follow-up discussion. The SPC provides a standardized method for rating teachers' instructional practices; however, coaches also take extensive field notes about teacher and student interactions (e.g., turn taking, questioning patterns, etc.). For stage three, the coach and teacher hold a 30-min debrief, comparing the lesson as designed to the lesson delivered, reflecting on its strengths and areas for improvement. The SPC, used regularly in the planning, observing, and debriefing stages, reinforce the performance targets for coaching. Teacher goals for improvement, however, are always dependent on a combination of teacher and student readiness.

### 3.4. Data analysis

Data analysis occurred in three steps. First, frequencies, means, and standard deviations were calculated for all independent and dependent variables. Second, multiple one-way repeated measure ANOVAs were conducted to reveal patterns of development for each of the Five Standards. In order to understand differences in teachers' ultimate level of implementation, a one-way ANOVA was conducted for group, where participants were assigned to a high ($n = 9$; $>8$) or low ($n = 12$; $\leq 8$) group based on their baseline observation. This analysis revealed if patterns of ultimate development varied by teachers' initial capacity—or starting points for learning—and reveals the extent to which the coach–teacher (expert–novice in the instructional model) collaboration assists development. Analyses examined how each coaching cycle influenced (a) Total Score, (b) individual standard scores, and (c) the number of activity centers in use. Mauchly's test of sphericity resulted in corrections as needed. In general, the Wilks' Lambda value, $F$ statistic, and a partial eta-squared value are reported. Effect sizes are defined (Cohen, 1988) as small ($<.20$), medium ($>.20$ and $<.79$) and large ($>.80$). Tests of within-subjects contrasts identify significant linear and quadratic trends in the data across cycles.

Finally, post-hoc pairwise comparisons, using the least significant difference (LSD) method, identified significant differences ($p < .05$) in mean performance across coaching cycles. Marked line graphs of development are provided. The results are reported by research questions (RQ) in three sections.

## 4. Results

### 4.1. Use of the Five Standards

RQ1 asked if teachers increased their use of individual standards across coaching cycles. Table 1 presents the means and standard deviations for each standard and Total Score by coaching cycle. Teacher use of each standard consistently increased from coaching cycle one to four. For Contextualization, Challenging Activities, and Total Score, the mean increased into cycle five as well. Joint Productivity and Instructional Conversation each showed a plateau effect during cycles five and six while Language/Literacy showed a plateau during cycles four and five. Only Contextualization and Total Score showed a decline during both cycles six and seven from a peak in cycle five. Standard deviations (SDs) increased, showing more teacher variation with each coaching cycle, from cycle one to three. The Instructional Conversation showed the greatest variation among teachers across all coaching cycles, with the exception of Contextualization in cycle seven.

The one-way repeated measure ANOVAs revealed that teacher growth was statistically significant in use of the Five Standards

**Table 1**
Means and standard deviations for five standards 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 | 1.81 | 2.33 | 3.14 | 3.57 | 3.48 | 3.48 | 3.38 |
| | SD | .81 | .86 | 1.06 | .75 | .98 | .87 | .92 |
| Language/Literacy Development (LLD) | M | 2.05 | 2.29 | 3.14 | 3.57 | 3.57 | 3.24 | 3.48 |
| | SD | .67 | .85 | 1.01 | .75 | .93 | 1.00 | .87 |
| Contextualization (CTX) | M | 1.43 | 2.05 | 2.67 | 3.19 | 3.29 | 3.14 | 2.76 |
| | SD | .68 | 1.07 | 1.46 | .98 | 1.23 | 1.11 | 1.26 |
| Challenging Activities (CA) | M | 1.62 | 2.10 | 2.76 | 3.14 | 3.29 | 3.24 | 3.29 |
| | SD | .50 | .70 | 1.14 | 1.01 | 1.06 | 1.14 | .96 |
| Instructional Conversation (IC) | M | 1.05 | 1.55 | 2.20 | 2.90 | 3.00 | 3.00 | 2.90 |
| | SD | .95 | 1.15 | 1.67 | 1.21 | 1.38 | 1.34 | 1.17 |
| Total Score | M | 8.00 | 10.24 | 13.90 | 16.24 | 16.62 | 16.10 | 15.81 |
| | SD | 2.51 | 4.11 | 6.03 | 4.00 | 5.00 | 4.58 | 4.12 |

Total N = 21.

individually and for Total Score by coaching cycle with large effect sizes: (a) Joint Productivity Wilks' Lambda = .13, $F(6, 15) = 17.54$, $p < .001$, partial eta-squared .88; (b) Language/Literacy Wilks' Lambda = .15, $F(6,15) = 14.29$, $p < .001$, partial eta-squared .85; (c) Contextualization Wilks' Lambda = .14, $F(6,15) = 15.15$, $p < .001$, partial eta-squared .86; (d) Challenging Activities Wilks' Lambda = .15, $F(6,15) = 14.76$, $p < .001$, partial eta-squared .86; (e) Instructional Conversation Wilks' Lambda = .21, $F(6,14) = 8.94$, $p < .001$, partial eta-squared .79; (f) Total Score Wilks' Lambda = .11, $F(6,15) = 20.87$, $p < .001$, partial eta-squared .89.

The LSD comparisons revealed significantly greater mean use of each standard at coaching cycle seven than at coaching cycle one: Joint Productivity from 1.81 to 3.38; Language/Literacy from 2.05 to 3.48; Contextualization from 1.43 to 2.76; Challenging Activities from 1.62 to 3.29; Instructional Conversation from 1.05 to 2.90; and Total Score from 8.00 to 15.81. In terms of overall fidelity to the model, the Total Score means indicate that performance-based instructional coaching accomplished teacher use of three standards simultaneously but not more. On average teachers achieved an enacting level of fidelity (i.e., $M = 12.50 < 17.49$) by coaching cycle seven ($M = 15.81$; SD = 4.12) rather than the highest integrating level of fidelity ($M > 17.50$).

### 4.2. Patterns of development

RQ2 investigated the nature of cycle-to-cycle patterns of development in more depth using repeated measures contrasts and graph line plots of means. The data revealed a significant linear trend for each individual standard and Total Score with medium—large effect sizes: Joint Productivity $F(1, 20) = 53.44$, $p < .001$ (partial eta-squared = .73); Language/Literacy $F(1, 20) = 40.27$, $p < .001$ (partial eta-squared = .67); Contextualization $F(1, 20) = 46.02$, $p < .001$ (partial eta-squared = .70); Challenging Activities $F(1, 20) = 54.82$, $p < .001$ (partial eta-squared = .73); Instructional Conversation $F(1, 20) = 63.53$, $p < .001$ (partial eta-squared = .77); Total Score $F(1, 20) = 74.97$, $p < .001$ (partial eta-squared = .79). This linear trend demonstrates that teachers consistently improved from cycles one to five.

The data also revealed a significant quadratic—or single bend—trend in the data for each standard and Total Score with a medium effect size: Joint Productivity $F(1, 20) = 30.92$, $p < .001$ (partial eta-squared = .61); Language/Literacy $F(1, 20) = 20.05$, $p < .001$ (partial eta-squared = .50); Contextualization $F(1, 20) = 22.73$, $p < .001$ (partial eta-squared = .53); Challenging Activities $F(1, 20) = 10.97$, $p = .003$ (partial eta-squared = .35); Instructional Conversation $F(1, 20) = 26.87$, $p < .001$ (partial eta-squared = .59); Total Score $F(1, 20) = 30.13$, $p < .001$ (partial eta-squared = .60). The quadratic trend demonstrates that teachers also experienced slight declines in development during coaching cycles six and seven.

Fig. 3 visually presents these significant linear and quadratic patterns of development comparing each standard to each other. Three broad patterns of development are revealed. First, from the first to last cycle of coaching, teachers used Joint Productivity and Language/Literacy at higher levels than the other standards. Joint Productivity and Language/Literacy appear to be the easiest standards for teachers to implement. Second, the increasing trajectory of Challenging Activities and Contextualization was similar up to cycle five, when Contextualization declined in use more than any other standard during coaching cycles six and seven. Finally, the Instructional Conversation was the least observed standard at the first coaching cycle and clearly the most difficult for teachers to implement. While teachers worked with a small group of students (i.e., had a teacher center), they did not achieve the highest level of enacting an Instructional Conversation as defined by the SPC.
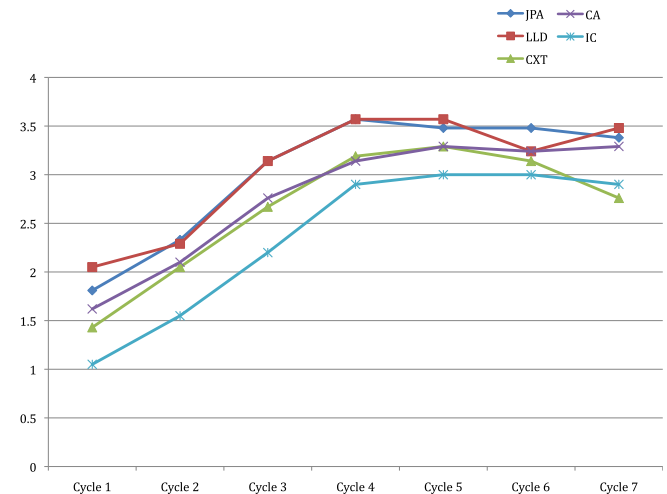


**Fig. 3.** Marked line graph of each standard in contrast to each other.

Least significant difference post-hoc comparisons specifically identify where significant teacher growth occurred across coaching cycles. Appendix A contains Tables A1–A6, with the mean differences, standard errors, and 95% confidence intervals for statistically significant ($p < .05$) cycle-to-cycle pairwise comparisons in teacher performance by individual standard and Total Score.

Five dominant patterns of development are noteworthy. First, Joint Productivity, Contextualization, and Challenging Activities showed significant mean difference growth between cycle one and each subsequent cycle and between cycle two and each subsequent coaching cycle (Tables A1, A3, and A4). Second, teachers did not experience significant growth in use of Language/Literacy and Instructional Conversation between coaching cycles one and two although significant growth occurred for each subsequent coaching cycle (Tables A2 and A5). Teachers were already using Language/Literacy to a higher degree than the other standards. In the case of Instructional Conversation, a vast majority of teachers did not use a teacher center between coaching cycles one and two; their focus was on implementation of independent centers.

Third, there was significant growth by teachers in the use of Challenging Activities and Instructional Conversations in later cycles of coaching (Tables A4 and A5). For Challenging Activities, teachers experienced significant growth between coaching cycles three and five. There was significant growth in Instructional Conversation use between cycle three and cycles five, six, and seven. The Instructional Conversation is the only standard where teachers made significant growth in each of the final coaching sessions.

Fourth, in rank ordering, teachers experienced the most change (mean difference) from coaching cycle one to seven in use of Instructional Conversation (1.85), followed in order by Challenging Activities (1.67), Joint Productivity (1.57), Language/Literacy (1.43), and Contextualization (1.33). Although the Instructional Conversation was the most difficult to implement, teachers experienced the most growth in use of this standard.

Fifth, Total Score mean differences revealed significant growth from coaching cycle one and coaching cycle two to each subsequent coaching cycle (Table A6). Total Score growth was also significant between coaching cycles three and five, with no other significant difference achieved beyond coaching cycle five. The largest increases in teacher growth occurred across coaching cycles one and four, as revealed in these cycle-to-cycle mean differences: one to two ($M = 2.24$), two to three ($M = 3.67$), three to four ($M = 2.33$), four to five ($M = .38$), five to six ($-.52$), and six to seven ($-.29$).

Table 2 presents the means and standard deviations for the Five Standards and coaching cycle by teachers in the high and low groups. Four patterns stand out: (a) The low group consistently enacted the Five Standards Instructional Model at a lower level than the high group across all coaching cycles as measured by Total Score; (b) only for Contextualization in cycles 4 and 6 does the lower group mean rise above the high group for any standard; (c) the low group teachers consistently enact the Instructional Conversation least; and (d) the low group most consistently struggled to enact Instructional Conversation, Contextualization, and Challenging Activities to a high level. These patterns demonstrate that teachers in the high and low groups required ongoing assistance unique to their development.

Fig. 4 presents a graph comparing high and low groups by Total Score across coaching cycles. One-way ANOVAs revealed significant high and low group differences at the time of baseline observations for each standard and Total Score: Joint Productivity $F(1, 20) = 5.71$, $p = .03$; Language/Literacy $F(1, 20) = 7.43$, $p = .01$; Contextualization $F(1, 20) = 12.07$, $p = .002$; Challenging Activities $F(1, 20) = 18.41$, $p < .001$; Instructional Conversation $F(1, 20) = 8.99$, $p = .007$; Total Score: $F(1, 20) = 33.94$, $p < .001$. Group differences were not significant ($p < .05$) for the seventh coaching cycle for Total Score or individual standards: Joint Productivity $F(1, 20) = 3.61$, $p = .07$; Language/Literacy $F(1, 20) = 2.72$, $p = .12$; Contextualization $F(1, 20) = .38$, $p = .54$; Challenging Activities $F(1, 20) = 2.01$, $p = .17$; Instructional Conversation $F(1, 20) = 3.82$, $p = .07$; and (c) Total Score $F(1, 20) = 3.58$, $p = .07$. These findings demonstrate that target-based instructional coaching effectively closes the gap between high and low group implementers through tailored coaching assistance within a teacher's zone of proximal development.

### 4.3. Teacher use of small group classroom organization

RQ3 asked whether coached teachers increased use of small group activity centers. The mean number of activity centers significantly increased from coaching cycle one ($M = 3.05$; $SD = 2.25$) to cycle seven ($M = 4.95$; $SD = 1.56$), with a small effect size (partial eta-squared = .21) for coaching cycle: Lower-bound $F(1, 20) = 5.15$, $p = .03$. No high/low group differences were detected: $F(1, 20) = .03$, $p = .88$.

During coaching cycle one, 35% of teachers used a whole class organization, 42.5% used small group activity centers, but only 22.5% included a teacher's instructional conversation as one of their activity centers. By coaching cycle seven, 91% of the teachers were using the teacher-led instructional conversation center; however, only 59% of teachers enacted that instructional conversation at the

highest level of quality. Overall, instructional coaching increased teachers' use of small group activity centers in number and quality.

## 5. Discussion

Knight and Wiseman (2005) called for evidence on how to support teachers in acquiring instructional skills for diverse students. This study focused on a performance-based instructional coaching model targeting instructional practices for diverse learners. Following a brief summary of significant findings, this performance-based instructional coaching model is discussed in light of its theoretical, practical, and research implications.

### 5.1. Significant findings

This descriptive study documents that performance-based instructional coaching, focused on use of the Five Standards Instructional Model, resulted in statistically significant teacher growth across seven cycles of coaching (RQ1) as measured by individual standards and Total Score. Teachers achieved an enacting level (i.e., $M = 12.50 < 17.49$), rather than an integrating ($M > 17.50$) level, of fidelity by coaching cycle seven. Overall, instructional coaching led to significant transfer of new teaching skills from a workshop to the classroom similar to studies on peer coaching (i.e., Showers, 1982; Showers & Joyce, 1996).

Teacher growth was also patterned across coaching cycles (RQ2). A linear trend in teacher growth demonstrated consistent improvement for teachers from cycles one to five. Teachers experienced a plateau or slight decline in development during coaching cycles six and seven. Some individual standards were more challenging to enact than others. For example, the Instructional Conversation yielded the most growth yet remained the lowest implemented standard. Teachers grew the least consistently in use of Contextualization. Teachers quickly focused on Language/Literacy Development and Joint Productivity (i.e., cycles one and two) while Challenging Activities and the Instructional Conversation were the focus of later coaching cycles (i.e., five to seven). Overall, teachers grew the most between coaching cycles one and four.

Studying the pattern of implementation among teachers whose starting points for learning were initially high or low revealed more nuanced patterns of development. The low group consistently enacted the Five Standards Instructional Model at a lower level than the high group across all coaching cycles as measured by Total Score. While teachers in both groups increased use of small group configurations (RQ3), teachers in the low group consistently

**Table 2**
Means and standard deviations for five standards by coaching cycle.

| Five Pedagogical Standards | | Cycle 1 | | Cycle 2 | | Cycle 3 | | Cycle 4 | | Cycle 5 | | Cycle 6 | | Cycle 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low |
| Joint Productive Activity (JPA) | M | 2.22 | 1.46 | 2.44 | 2.25 | 3.33 | 2.92 | 3.78 | 3.31 | 3.78 | 3.23 | 3.56 | 3.38 | 3.78 | 3.08 |
| | SD | .83 | .66 | 1.01 | .75 | 1.00 | 1.12 | .67 | .85 | .67 | 1.09 | .88 | .87 | .67 | .95 |
| Language/Literacy Development (LLD) | M | 2.44 | 1.77 | 2.56 | 2.08 | 3.33 | 2.92 | 3.78 | 3.31 | 3.78 | 3.38 | 3.56 | 2.92 | 3.78 | 3.15 |
| | SD | .73 | .44 | .88 | .79 | 1.00 | 1.04 | .67 | .85 | .67 | 1.04 | .88 | 1.04 | .67 | .99 |
| Contextualization (CTX) | M | 1.89 | 1.08 | 2.33 | 1.83 | 3.00 | 2.46 | 3.00 | 3.31 | 3.33 | 3.15 | 3.22 | 3.00 | 2.89 | 2.54 |
| | SD | .78 | .28 | 1.22 | .94 | 1.50 | 1.39 | 1.12 | .85 | 1.32 | 1.21 | 1.20 | 1.08 | 1.36 | 1.27 |
| Challenging Activities (CA) | M | 2.00 | 1.31 | 2.33 | 1.92 | 2.89 | 2.62 | 3.56 | 2.69 | 3.56 | 3.00 | 3.33 | 3.08 | 3.56 | 2.92 |
| | SD | .00 | .48 | 1.00 | .29 | 1.05 | 1.19 | .88 | 1.11 | .88 | 1.15 | 1.00 | 1.26 | .88 | 1.12 |
| Instructional Conversation (IC) | M | 1.67 | .62 | 1.89 | 1.17 | 2.44 | 1.77 | 3.00 | 2.46 | 3.33 | 2.46 | 3.50 | 2.62 | 3.33 | 2.38 |
| | SD | 1.00 | .65 | 1.36 | .94 | 1.59 | 1.74 | 1.22 | 1.45 | 1.41 | 1.45 | .93 | 1.45 | 1.00 | 1.19 |
| Total Score | M | 10.22 | 6.23 | 11.56 | 9.25 | 15.00 | 12.85 | 17.11 | 15.00 | 17.89 | 15.38 | 17.00 | 15.08 | 17.44 | 14.08 |
| | SD | 1.92 | 1.30 | 5.10 | 3.05 | 5.74 | 6.16 | 3.59 | 4.64 | 4.20 | 5.36 | 4.36 | 4.77 | 3.36 | 4.54 |

High group $n = 9$; low group $n = 12$.
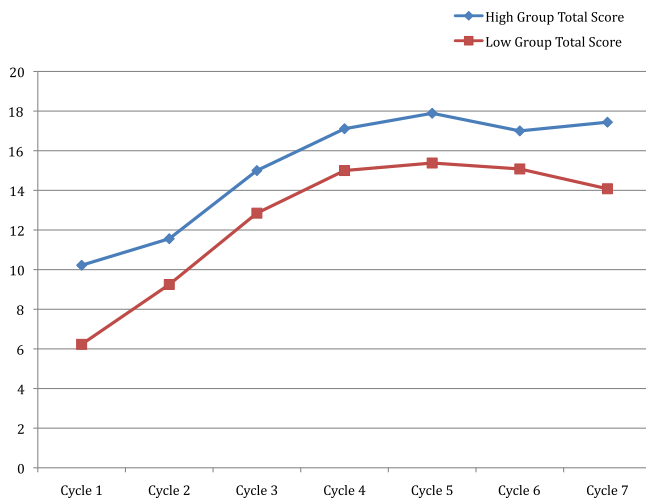
**Fig. 4.** Marked line graph of total score by high and low group.

struggled to enact the Instructional Conversation, Contextualization, and Challenging Activities to a high level in those small groups.

These patterns and differences between high and low group implementers demonstrate that teachers benefit from ongoing assistance unique to their needs. Most importantly, these findings demonstrated that target-based instructional coaching, when tailored to teachers' needs, is able to statistically close the pedagogical gap between teachers in the high and low groups over time.

### 5.2. Theoretical implications

These findings make four main contributions to the emerging theoretical knowledge base on instructional coaching for diversity. First, the efficacy of coaching in general (Cornett & Knight, 2009; Joyce & Showers, 1995; Knight, 2009a; Sparks & Hirsh, 1997; Speck & Knipe, 2001) and instructional coaching in particular are supported (Knight, 2009b). This study demonstrates that as a follow up to a 30-h workshop, instructional coaching had a significant, positive, and generally linear impact on teacher growth across seven coaching cycles. Instructional coaching, therefore, leads to significant teacher change, which is valuable in an international era demanding evidence of greater teacher effectiveness (e.g., Cobbold, 2010; The Sutton Trust, 2010; Wang, Coleman, Coley, & Phelps, 2003). Such findings provide "existence proof" (Borko, 2004, p. 5) of teacher change in an area that has received limited attention to date.

Second, this study demonstrates the value of combining instructional coaching with concrete and sociocultural performance targets. In particular, the SPC provided shared and actionable content anchoring the coach–teacher partnership as well as a valid and reliable tool for studying its effectiveness. Cornett and Knight (2009, p. 211) argue that use of valid coaching tools, such as the SPC, are important in understanding and evaluating best practices for coaching.

Third, knowledge about practice in the coaching process was collaboratively co-constructed, allowing a novice to work "closely with an expert in joint problem solving in the zone of proximal development" (Rogoff, 1990, p. 141). Coaching created opportunities for ongoing cycles of assessment and assistance to inform teacher improvement. This model not only requires coaches themselves to possess requisite, deep, and nuanced understandings of coaching targets but also requires dispositional attributes that respect teacher readiness to incorporate performance targets into their practice. While teacher performance led development, the coaching process accelerated growth.

Fourth, to improve performance-based instructional coaching, the pattern of teacher change was just as important to understand as the amount of change measured itself. While teachers significantly increased use of the Five Standards Instructional Model with seven coaching sessions, all coached teachers were not able to reach the highest level of fidelity to the model. Some standards in the model were more difficult to learn for teachers initially identified as low implementers. Patterns of development underscore the importance of differentiating professional development in light of observable teacher needs. These results may also substantiate other findings from multicultural research. For example, teachers who struggle with developing cognitively challenging activities may struggle with raising their learning expectations for diverse students. An inability to enact a high-level Instructional Conversation may be rooted in a fear of what greater student autonomy means for classroom management. The data on Contextualization indicate that teachers only made incidental—rather than deep, integrated—connections to students' knowledge from home, school, or community when teaching academic content. Teachers consistently underestimate, therefore, the importance of contextualizing new learning in their students' previous knowledge. Although this may be a result of district pacing guides, it may also indicate a lack awareness of students' cultural communities and/or the dominance of teachers' own cultural "frames of references" (Sleeter, 2008, p. 561).

The question is whether more coaching cycles alone can overcome barriers to implementing these more challenging standards to a high level. Sleeter (2008, p. 561) contends that instructional practices are only one part of learning to "teach *better* and *more equitably*." This would suggest that instructional coaching would best be part of a comprehensive professional development program simultaneously exploring teacher knowledge, dispositions, and practices for teaching diverse learners.

### 5.3. Practical implications for improving professional development

Three implications stand out for improving the professional development experience—30-h workshop plus coaching sessions—for teachers focused on the Five Standards Instructional Model. First, both the workshop and coaching process should highlight the pattern of incorporating the standards into teaching. Teachers are most prepared to require greater sustained language and literacy use by students, including a shift to more negotiated, collaborative, and challenging activities. These same standards—Language/Literacy, Joint Productivity, and Challenging Activities—were implemented to the same degree and order for cycle seven (See Fig. 3). This suggests some value in aligning the coaching process and emphasis to the revealed pattern of teacher change. Teachers could be encouraged to initially focus on Language/Literacy, Joint Productivity, and Challenging Activities, and then turn their focus to Contextualization and the Instructional Conversation exclusively during the later stages of coaching.

Second, emphasis on the instructional conversation should be given greater emphasis in the professional development. During the 30-h workshop, the Instructional Conversation is the focus of the final of five days of learning, when teachers are likely overwhelmed with new learning. Coaching outcomes could be improved by offering the Instructional Conversation day at a separate time and as a single day, later in the coaching process, to match teacher implementation patterns.

Third, the emphasis given to Contextualization presents a unique challenge for professional development. Teachers show the least growth and are least able to sustain growth with this

standard. Yet, research suggests strong benefits for diverse learners when knowledge from home, school, and community is the basis and connection for new learning (e.g., González, Moll, & Amanti, 2005). Further investigation is needed to understand if the quadratic trend for Contextualization is an artifact of the professional development, systemic to development, a by-product of high stakes testing pressures, district pacing guides, or teacher's unexamined beliefs.

### 5.4. Research implications

This study contributes to the coaching research base in several ways. It confirms performance-based instructional coaching as an effective professional development strategy for eliciting teacher change and increasing skill transfer (Cornett & Knight, 2009; Speck & Knipe, 2001). This study improved on previous instructional coaching studies, which relied mostly on self-report data (Knight, 2009c), by defining performance targets with a valid and reliable classroom observation rubric. The outcomes also document the value of understanding the coach—teacher partnership in terms of what it means to work collaboratively within a teacher's zone of proximal development.

Four implications for future research are evident. First, qualitative investigation of coaching conversations and notes would uncover deeper understandings of teacher change and the challenges associated with implementing use of the Five Standards Instructional Model. In essence, the coaching conversations would reveal unique and common obstacles to progress in the zone of proximal development. Second, further research is needed to investigate how many coaching cycles across what period of time result in the best teacher achievement. Achievement, of course, does not equate with sustainability. This distinction between achievement and sustainability deserves further investigation. Third, additional longitudinal studies must confirm the veracity of the observed linear and quadratic developmental trends. Fourth, future studies should investigate the relationship between coached teacher performance and student achievement. While several studies (e.g., Doherty & Hilberg, 2007) have established a significant relationship between teacher's natural use of the Five Standards Instructional Model and student achievement, such results have not yet been replicated in the context of this instructional coaching model.

With international interest and experimentation with performance pay for K-12 teachers increasing (Podgorksy & Springer, 2007), rigorous research establishing valid, reliable, and value-added models and measures of professional development are of growing importance. This study suggests that measurable performance targets, when mutually viewed as valid, reliable, and valued, provided the potential for greater accountability for instructional coaching outcomes at the teacher, coach, and school/district levels. At the very least, incorporating measurable performance targets avoids the weakness of other instructional coaching research (Cornett & Knight, 2009) by providing reliable evidence that teachers are implementing new teacher practices.

### 5.6. Conclusion

Identifying the content and outcomes of effective professional development for culturally and linguistically diverse students is an urgent need for countries with increasingly underachieving diverse student populations. Teachers benefit from knowing what pedagogy and classroom organizations promote achievement among diverse students. Performance-based instructional coaching, defined from a sociocultural perspective, like the Five Standards Instructional Model itself, is an important development. This study demonstrates both the efficacy and promise of performance-based instructional coaching as a value-added professional development strategy.

### Appendix. Supplementary material

Supplementary material related to this article can be found at doi:10.1016/j.tate.2010.11.006.

### References

Barmby, P. (2006). Improving teacher recruitment and retention: the importance of workload and pupil behaviour. *Educational Research, 48*(3), 247—265.

Bernhard, J. K., Lefebvre, M. L., Chud, G., & Lange, R. (1997). The preparation of early childhood educators in three Canadian areas of immigrant influx: diversity issues. *Canadian Children, 22*(1), 26—34.

Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher, 33*(8), 3—15.

Cobbold, T. (2007, September 7). *Addressing the social divide in education.* Retrieved from Save Our Schools: Fighting for Equity in Education website. http://www.saveourschools.com.au.

Cobbold, T. (2010, August 2). *Reducing the achievement gap between rich and poor is a national priority.* Retrieved from Save Our Schools: Fighting for Equity in Education website. http://www.saveourschools.com.au.

Cochran-Smith, M., & Zeichner, K. M. (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Erlbaum Associates.

Cornett, J., & Knight, J. (2009). Research on coaching. In J. Knight (Ed.), *Coaching: Approaches and perspectives* (pp. 192—216). Thousand Oaks, CA: Corwin Press, Inc.

Crane, T. G. (2002). *The heart of coaching: Using transformational coaching to create a high-performance culture* (2nd ed.). San Diego, CA: FTA Press.

Dalton, S. S. (2008). *Five standards for effective teaching: How to succeed with all learners, grades K-8.* San Francisco, CA: John Wiley & Sons, Inc.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Educational Policy Analysis Archive, 8*(1). Retrieved from. http://epaa.asu.edu/ojs/article/view/392.

Darling-Hammond, L., & Sykes, G. (1999). *Teaching as the learning profession: Handbook of policy and practice.* San Francisco: Jossey-Bass Publishers.

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., Pinal, A., & Tharp, R. G. (2003). Five standards and student achievement. *NABE Journal of Research and Practice, 1*(1), 1—24.

Estrada, P. (2005). The courage to grow: a researcher and teacher linking professional development with small-group reading instructional and reading achievement. *Research in the Teaching of English, 39*(1), 320—364.

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.

Freire, P. (1994). *Pedagogy of the oppressed*. New York: Continuum.

Figueroa, P. (2004). Multicultural education in the United Kingdom: historical development and current status. In J. A. Banks, & C. A. McGee Banks (Eds.), *Handbook of research on multicultural education* (2nd ed.). (pp. 997–1026) San Francisco, CA: Jossey-Bass.

Garret, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

González, N., Moll, L. C., & Amanti, C. (2005). *Funds of knowledge: Theorizing practice in households, communities, and classrooms*. New York: Routledge.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Haworth, P. (2009). The quest for a mainstream EAL pedagogy. *Teacher College Record, 111*(9), 2179–2208.

Hawley, W., & Valli, L. (1999). The essentials of effective professional development: a new consensus. In L. Darling-Hammond, & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 127–150). San Francisco: Jossey-Bass Publishers.

Hilberg, R. S., Chang, J. M., & Epaloose, G. (2003). *Designing effective activity centers for diverse learners*. Santa Cruz, CA: Center for Research on Education, Diversity & Excellence, University of California, Santa Cruz.

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, New York: 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.

Hilberg, R. S., Tharp, R. G., & DeGeest, L. (2000). The efficacy of CREDE's standards-based instruction in American Indian mathematics classes. *Equity and Excellence in Education, 33*(2), 32–39.

Hill, B., & Allan, R. (2004). Multicultural education in Australia: historical development and current status. In J. A. Banks, & C. A. McGee Banks (Eds.), *Handbook of research on multicultural education* (2nd ed.). (pp. 979–996) San Francisco, CA: Jossey-Bass.

Hollins, E., & Guzman, M. T. (2005). Research on preparing teachers for diverse populations. In M. Cochran-Smith, & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 477–548). Mahwah, NJ: Lawrence Erlbaum.

Ingersoll, R. M. (2004). *Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers? Report prepared for renewing our schools: Securing our future, a National task force on public education*. Washington, DC: Center for American Progress and Institute for America's Future.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development*. White Plains, NY: Longman.

Knight, J. (2004). Instructional coaches make progress through partnership. *Journal of Staff Development, 25*(2), 32–37.

Knight, J. (2009a). *Coaching: Approaches and perspectives*. Thousand Oakes, CA: Corwin Press.

Knight, J. (2009b). Introduction. In J. Knight (Ed.), *Coaching Approaches and perspectives* (pp. 1–6). Thousand Oakes, CA: Corwin Press.

Knight, J. (2009c). Instructional coaching. In J. Knight (Ed.), *Coaching Approaches and perspectives* (pp. 29–55). Thousand Oakes, CA: Corwin Press.

Knight, S. L., & Wiseman, D. L. (2005). Professional development for teachers of diverse students: a summary of the research. *Journal of Education for Students Placed at Risk, 10*(4), 387–405.

Kowal, J., & Steiner, L. (2007). *Instructional coaching*. Retrieved November 6, 2009, from. www.centerforcsri.org.

Lewis, C., Enciso, P., & Moje, E. B. (2007). *Reframing sociocultural research on literacy: Identify, agency, and power*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Mackel, C. (2002). *An ethic of care: Effective programs for beginning teachers. Report for the commonwealth department for education, science, and training*. Retrieved from. http://www.dest.gov.au/schools/publications/2003/beginningteachers.pdf.

National Center for Education Statistics.. (2010). *The Condition of education 2010*. Washington, DC: Institute of Educational Sciences, U.S. Department of Education. Retrieved from. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010028.

Podgorksy, M. J., & Springer, M. G. (2007). Teacher performance pay: a review. *Journal of Policy Analysis and Management, 26*(4), 909–949.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Rogoff, B., & Wertsch, J. V. (1984). *Children's learning in the zone of proximal development*. San Francisco: Jossey-Bass.

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.

Sherris, A. (2010, July). *Coaching language teachers. CALdigest*. Washington, DC: Center for Applied Linguistics. Retrieved from. http://www.cal.org/resources/digest/index.html.

Showers, B. (1982). *Transfer of training: the contribution of coaching*. Eugene, OR: Center for Educational Policy and Management. Retrieved from Education Resources Information Center website. http: www.eric.ed.gov.

Showers, B., & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership, 53*(6), 12–16.

Skerrett, A. (2008). "Going the race way": biographical influences on multicultural and antiracist English curriculum practices. *Teaching and Teacher Education, 24*(7), 1813–1826.

Sleeter, C. E. (2008). Preparing White teachers for diverse students. In M. Cochran-Smith, S. Feiman-Nemser, D. J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed.). (pp. 559–582) New York: Routledge, Taylor & Francis Group and the Association of Teacher Education.

Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.

Speck, M., & Knipe, C. (2001) *Why can't we get it right? Professional development in our schools*. Thousand Oaks, CA: Corwin Press.

Teemant, A., Tyra, S., & Wink, J. (2009). Evaluating the effectiveness of the CREDE coaching model. In *Proceedings from the 3rd Annual center for research on education, diversity & excellence conference*. Berkeley, San Francisco, CA: University of California.

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: Cambridge University Press.

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

The Sutton Trust. (2010, April). *Education mobility in England: The link between the education levels of parents and the educational outcomes of teenagers*. Retrieved from. http://www.suttontrust.com/annualreports.asp.

U.S.Department of Education. (2009, November). *Race to the top executive summary*. Retrieved from. http://www2.ed.gov/programs/racetothetop/index.html.

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

Vygotsky, L. S. (1997). *Educational psychology*. Boca Rotan, FL: CRC Press.

Wang, A. H., Coleman, A. B., Coley, R. J., & Phelps, R. P. (2003, May). *Preparing teachers around the world*. Princeton, NJ: Educational Testing Service. Retrieved from. http://www.ets.org/research/policy_research_reports/pic-preteach.

Wink, J., & Putney, L. (2002). *Vision of Vygotsky*. Boston, MA: Allyn & Bacon.

Zepeda, S. J. (2008). *Professional development: What works*. Larchmont, NY: Eye on Education.